

Financial Valuation of Mortality Risk via the Instantaneous Sharpe Ratio: Applications to Pricing Pure Endowments

Version: 4 November 2005

Moshe A. Milevsky

Schulich School of Business
York University
Toronto, Ontario, M3J 1P3
milevsky@yorku.ca

S. David Promislow

Department of Mathematics and Statistics
York University
Toronto, Ontario, M3J 1P3
promis@yorku.ca

Virginia R. Young

Department of Mathematics
University of Michigan
Ann Arbor, Michigan, 48109
vryoung@umich.edu

Abstract: We develop a theory for pricing non-diversifiable mortality risk in an incomplete market. We do this by assuming that the company issuing a mortality-contingent claim requires compensation for this risk in the form of a pre-specified instantaneous Sharpe ratio. We prove that our ensuing valuation formula satisfies a number of desirable properties. For example, we show that it is subadditive in the number of contracts sold. A key result is that if the hazard rate is stochastic, then the risk-adjusted survival probability is greater than the physical survival probability, even as the number of contracts approaches infinity.

Keywords: Stochastic mortality, pricing, Sharpe ratio, non-linear partial differential equations.

JEL Classification: G13; G22; C60.

MSC 2000: 91B30; 91B70.

1. Introduction

Insurance markets are incomplete for several reasons, and in this paper we focus on the incompleteness arising from two sources: (1) The timing of insurance payments is generally determined by a jump process – such as, when individuals die or when cars crash, and (2) insurers cannot buy and sell insurance contracts frictionlessly, if at all. Actuaries traditionally assume that one can eliminate the uncertainty associated with the jump process by selling a large number of insurance contracts. By invoking the law of large numbers, actuaries thereby replace the random occurrence of insurance payments with a deterministic schedule.

This assumption is the reason that Brennan and Schwartz (1976), or more recently Boyle and Hardy (2003), were able to extend the Black-Scholes formula to price a derivative instrument with a random maturity date. They assumed that the issuer sold a sufficiently large number of policies so that the issuer only had to contend with a deterministic schedule of maturing derivatives. This simplifying technique permeates most of the literature that combines finance and insurance. While each individual’s payment occurs at a random time, one assumes that a large portfolio becomes deterministic and, thus, safe from a mortality point of view.

Using the language of modern portfolio theory, under this assumption, the idiosyncratic risk – that is, the standard deviation per contract – goes to zero if the insurer sells enough contracts, and so the risk is diversifiable. However, because the insurer can only sell a finite number of insurance policies, it is impossible for the insurer to eliminate the risk that the experience will differ from what is expected. The risk associated with selling a finite number of insurance contracts is what we call the *finite portfolio risk*.

In addition to the finite portfolio risk, there is a risk arising from the fact that the parameters underlying the claim occurrence process are random themselves. For example, for an insurance payment that is contingent on the survival or death of an individual, one models the claim occurrence process as a Poisson process with intensity λ , also called the hazard rate. When the individual dies, then the process jumps and is “killed” at that time. Now, if λ is a deterministic function of time, then the risk associated with selling a finite number of insurance contracts, each of whose claim occurrence is dictated by independent and identically distributed killed Poisson processes, is the finite portfolio risk. If λ is stochastic, then there is an additional risk that we call the *stochastic mortality risk*, a special case of stochastic parameter risk. Even as the insurer sells an arbitrarily large number of contracts, the systematic stochastic mortality risk remains because of the second source of incompleteness, namely, the inability to buy and sell insurance contracts

frictionlessly.

One could also define similar risks associated with the size of insurance payments. However, for concreteness in this paper, we focus on the rate of claim occurrence and leave considerations concerning the size of the payments to future research. In particular, we determine how to value the risks associated with a specific contract called a *pure endowment*. A pure endowment pays \$1 to an individual at a given time T if the individual survives to that time. Pure endowments are important in that they are the building blocks of life annuities. They are also analogous to defaultable zero-coupon bonds because of the uncertainty of the payment at time T .

We emphasize that stochastic parameter risk is different from model specification risk. For stochastic parameter risk, we assume that we have specified the model precisely, including the stochastic process that the model parameter follows. However, if one cannot buy and sell insurance contracts frictionlessly, one cannot eliminate the systematic risk inherent in this stochastic process. On the other hand, if an insurer were to issue a pure endowment contract to infinitely many individuals, each of whose independent time-of-death follows the same stochastic hazard rate, then the insurer could completely hedge the stochastic mortality risk by selling term life insurance to infinitely many individuals, each of whose independent mortality follows the same stochastic hazard rate as the group buying the pure endowment. This remark is the large-scale version of the statement that if one could sell both a pure endowment and a term life insurance contract to the same person (with the same maturity date and same mortality-contingent payment), then there would be no risk. In this case, the price for the combined contract would be the price for a default-free zero-coupon bond.

In this paper, we consider the case in which the issuer of a pure endowment *cannot* sell term life insurance to hedge the stochastic mortality risk. (In future work, we allow the issuer to hedge the risk partially by selling life insurance to an individual whose stochastic mortality is correlated with that of the purchaser of the pure endowment.) We argue that survival probabilities are uncertain and that this uncertainty is correlated across individuals in a population – mostly due to medical breakthroughs or environmental factors that affect the entire population. For example, if there is a positive probability that medical science will find a cure for cancer during the next thirty years, this will influence aggregate mortality patterns. The uncertainty regarding the evolution of survival probabilities induces a mortality dependence that cannot be diversified by selling more contracts. This risk induces a mortality risk premium that should be priced by the market and whose magnitude depends on a representative investor’s risk aversion or demanded compensation for risk.

We are not the first to recognize that mortality rates themselves should be viewed as stochastic. Biffis (2005), Schrager (2005), Dahl (2004), as well as Milevsky and Promislow (2001) and Soininen (1995), used and calibrated diffusion processes to model the force of mortality. Also, demographers and actuaries, such as Lee and Carter (1992), Olivieri (2001), and DiLorenzo and Sibillo (2003), developed methods for projecting mortality. While some practitioners, such as Smith, Moran, and Walczak (2003), and academics, such as Cox and Lin (2004) and Cairns, Blake, and Dowd (2004), implicitly recognize that mortality risk is being priced by the market, they did not derive the actual value of this risk from first principles. In related work, Blanchett-Scalliet, El Karoui, and Martellini (2005) value assets that mature at a random time by using the principle of no arbitrage; the resulting pricing rule is, therefore, linear. However, for insurance markets, one cannot assert that no arbitrage holds, so we take a different approach to valuing insurance contracts.

We value a pure endowment by assuming that the insurance company issuing the contract is compensated for risk via the so-called *instantaneous Sharpe ratio* of a suitably-defined portfolio. Specifically, we assume that the insurance company picks a target ratio of expected excess return to standard deviation, denoted by α , and then determines a price for a pure endowment that yields this pre-determined α for the corresponding portfolio. Our results do not depend on using a particular diffusion for the hazard rate. In future research, we plan to calibrate this model to a generalized mean-reverting process for mortality rates, similar to the work by Norberg (2004) in the context of interest rates.

Our methodology recovers a number of results that one expects within the context of insurance, but we also obtain new insights into the breakdown of traditional insurance pricing. For example, we prove that if the hazard rate is deterministic, then as the number of contracts approaches infinity, the price of a pure endowment collapses to the discounted expected payment using the physical probability measure to value the mortality risk, regardless of the target value of the Sharpe ratio. In other words, if the stochastic mortality risk is not present, then the price for a large portfolio reflects this and reduces to the “usual” expected value pricing rule in the limit.

A key result of the paper is that if the hazard rate is stochastic, then the financial value of the pure endowment is greater than the above-mentioned discounted expected payment, even as the number of contracts approaches infinity. Furthermore, our valuation operator is subadditive and satisfies a number of other appealing properties. Finally, our methodology allows us to decompose the value of any portfolio of pure endowment policies into a systematic component (due to uncertain aggregate mortality) and a non-systematic component (due to insuring a finite number of policies); see equation (4.55).

The remainder of this paper is organized as follows. In Section 2, we present our financial market, describe how to use the instantaneous Sharpe ratio to price the pure endowment, and derive the resulting partial differential equation that the price solves. In Section 3, we discuss qualitative properties of the risk-adjusted price from Section 2 and show that it shares many properties with the (static) standard deviation premium principle (Gerber, 1979). In Section 4, we study properties of the price for n conditionally independent and identically distributed pure endowment risks. In particular, in Theorem 4.11, we show that the price is subadditive with respect to n , and in Theorem 4.13, we show that the risk charge per person decreases as n increases. We also prove that if the hazard rate is deterministic, then the risk charge per person goes to zero as n goes to infinity (Theorem 4.20 and Corollary 4.21). Moreover, we prove that if the hazard rate is stochastic, then the risk charge per person is positive as n goes to infinity, which reflects the fact that the mortality risk is not diversifiable (Theorem 4.20 and Corollary 4.22). Section 5 concludes the paper.

2. Instantaneous Sharpe Ratio

In this section, we describe a pure endowment contract and present the financial market in which the issuer of this contract invests. We obtain the hedging strategy for the issuer of the pure endowment. We describe how to use the instantaneous Sharpe ratio to price the pure endowment and derive the resulting partial differential equation that the price solves. In Sections 3 and 4, we determine properties of the price.

2.1. Mortality Model and Financial Market

We begin with a stochastic model for mortality. We assume that the hazard rate λ (or force of mortality) of an individual follows a diffusion process such that if the process begins at $\lambda_0 > \underline{\lambda}$ for some positive constant $\underline{\lambda}$, then $\lambda_t > \underline{\lambda}$ for all $t \in [0, T]$. Thus, we require that the volatility of λ goes to zero as $\lambda \rightarrow \underline{\lambda}$ from the right, and we require that the drift of λ is positive for λ close to $\underline{\lambda}$. The reason for requiring λ to have a positive lower bound $\underline{\lambda}$ will be apparent later. From a modeling standpoint, $\underline{\lambda}$ could represent the hazard rate remaining (say, from accidents) after all biological causes of death have been removed.

Specifically, we assume that

$$d\lambda_t = a(\lambda_t, t)dt + b(t)(\lambda_t - \underline{\lambda})dW_t^\lambda, \quad (2.1)$$

in which W^λ is a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The volatility b is either identically zero, or it is a continuous function of time t bounded below by a

positive constant κ on $[0, T]$. The drift a is a Hölder continuous function of λ and t for which there exists $\epsilon > 0$ such that if $0 < \lambda - \underline{\lambda} < \epsilon$, then $a(\lambda, t) > 0$ for all $t \in [0, T]$. After Lemma 3.3 below, we add additional requirements for a . Note that if $b \equiv 0$, then λ is deterministic, and in this case, we write $\lambda(t)$ to denote the deterministic hazard rate at time t .

Suppose an insurer issues a pure endowment to an individual that pays 1 at time T if the individual is alive at that time. In Section 2.2, to determine the value of the pure endowment, we will create a portfolio composed of the obligation to pay this pure endowment and of default-free zero-coupon bonds that pay 1 at time T regardless of the state of the individual. Therefore, we require a model for bond prices, and we use a model based on the short rate and the bond market's price of risk.

The dynamics of the short rate r , which is the rate at which the money market increases, are given by

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dW_t, \quad (2.2)$$

in which μ and $\sigma \geq 0$ are deterministic functions of the short rate and time, and W is a standard Brownian motion with respect to the probability space $(\Omega, \mathcal{F}, \mathbf{P})$, independent of W^λ . We assume that μ and σ are such that $r \geq 0$ almost surely.

From the principle of no-arbitrage in the bond market, there is a market price of risk q for the bond that is adapted to the filtration generated by W ; see, for example, Lamberton and Lapeyre (1996) or Björk (1998). Moreover, the bond market's price of risk at time t is a deterministic function of the short rate and of time, that is, $q_t = q(r_t, t)$. Thus, the time t price of a T -bond is given by

$$F(r, t; T) = \mathbf{E}^{\mathbf{Q}} \left[e^{-\int_t^T r_s ds} \middle| r_t = r \right], \quad (2.3)$$

in which \mathbf{Q} is the probability measure with Radon-Nikodym derivative with respect to \mathbf{P} given by

$$\frac{d\mathbf{Q}}{d\mathbf{P}} = e^{-\int_0^T q(r_s, s)dW_s - \frac{1}{2} \int_0^T q^2(r_s, s)ds}. \quad (2.4)$$

It follows that $W^{\mathbf{Q}}$, with $W_t^{\mathbf{Q}} = W_t + \int_0^t q(r_s, s)ds$, is a standard Brownian motion with respect to \mathbf{Q} .

From Björk (1998), we know that the bond price F solves the following partial differential equation (pde):

$$F_t + \mu^Q(r, t)F_r + \frac{1}{2}\sigma^2(r, t)F_{rr} - rF = 0, \quad F(r, T; T) = 1, \quad (2.5)$$

in which $\mu^Q = \mu - q\sigma$. In this paper, the horizon T is fixed, so henceforth we drop T from the notation of F . We can use this pde to obtain the dynamics of the bond price $F(r_s, s)$, in which we think of $r_t = r$ as given and $t \leq s \leq T$. Indeed,

$$\begin{cases} dF(r_s, s) = (r_s F(r_s, s) + q(r_s, s)\sigma(r_s, s)F_r(r_s, s))ds + \sigma(r_s, s)F_r(r_s, s)dW_s, \\ F(r_t, t) = F(r, t). \end{cases} \quad (2.6)$$

As an aside, we could use other models commonly used in the literature and obtain the same conclusion that we reach after equation (2.17), namely, that we can factor the T -bond price from the “mortality price.” We use a model involving the short rate and the bond market’s price of risk for ease of presentation.

2.2. Pricing via the Instantaneous Sharpe Ratio

The insurer faces the unhedgeable risk that the individual’s living or dying will be different from expected; therefore, the insurer demands a return greater than the sum of the return r on the money market and the “return” λ from the mortality component. One measure of the risk that the insurer takes is the standard deviation of the change in the portfolio. A natural tie between the excess return and the standard deviation is the ratio of the former to the latter, the so-called *instantaneous Sharpe ratio*. In what follows, we find the hedging strategy to minimize the local variance of the change in the portfolio, then we set the price of the pure endowment so that the resulting instantaneous Sharpe ratio equals a given constant. We could set the instantaneous Sharpe ratio equal to a function of λ and t , but we choose a constant for simplicity.

The market for insurance is incomplete; therefore, there is no unique pricing mechanism. To value contracts in this market, one must assume something about how risk is “priced.” For example, one could use the principle of equivalent utility (see Zariphopoulou (2001) for a review) or the Esscher transform (Gerber and Shiu, 1994) to price the risk. We employ the instantaneous Sharpe ratio because of its analogy with the bond market’s price of risk and because of the desirable properties of the resulting price. Because of these properties, we anticipate that our pricing methodology will prove useful in pricing risks in other incomplete markets.

Denote the value (price) of the pure endowment by $P = P(r, \lambda, t)$, in which we explicitly recognize that the price of the pure endowment will depend on the short rate r and

the hazard rate λ at time t . Suppose the insurer creates a portfolio Π with value Π_t at time t . The portfolio contains the obligation to pay the pure endowment at time T if the individual is alive at that time, namely $-P$. Additionally, the insurer holds π_t T -bonds. Thus, $\Pi_t = -P(r_t, \lambda_t, t) + \pi_t F(r_t, t)$.

By Itô's Lemma (Protter, 1995), the value of the portfolio at time $t + h$ with $h > 0$, namely Π_{t+h} , equals

$$\begin{aligned} \Pi_{t+h} = & \Pi_t - \int_t^{t+h} \mathcal{D}^\mu P(r_s, \lambda_s, s) ds + \int_t^{t+h} \sigma(r_s, s) (\pi_s F_r(r_s, s) - P_r(r_s, \lambda_s, s)) dW_s \\ & - \int_t^{t+h} b(s) (\lambda_s - \underline{\lambda}) P_\lambda(r_s, \lambda_s, s) dW_s^\lambda + \int_t^{t+h} P(r_s, \lambda_s, s) (dN_s - \lambda_s ds) \\ & + \int_t^{t+h} \pi_s (r_s F(r_s, s) + q(r_s, s) \sigma(r_s, s) F_r(r_s, s)) ds, \end{aligned} \quad (2.7)$$

in which \mathcal{D}^m , with $m = m(r, t)$ a deterministic function of the short rate and time, is an operator defined on the set of appropriately differentiable functions on $\mathbf{R}^+ \times (\underline{\lambda}, \infty) \times [0, T]$ by

$$\mathcal{D}^m v = v_t + m v_r + \frac{1}{2} \sigma^2 v_{rr} + a v_\lambda + \frac{1}{2} b^2 (\lambda - \underline{\lambda})^2 v_{\lambda\lambda} - \lambda v. \quad (2.8)$$

Also, in (2.7), N denotes a Poisson process with stochastic parameter λ . Thus, Π jumps in value by P when an individual dies because the insurer is no longer responsible for paying 1 at time T .

In this single-life case, the process Π is “killed” when the individual dies. If we were to consider the price $P^{(n)}$ for n conditionally independent and identically distributed lives (conditionally independent given the hazard rate), then N would be a Poisson process with stochastic parameter $n\lambda$ such that Π jumps by $P^{(n)} - P^{(n-1)}$ when an individual dies. We consider $P^{(n)}$ later and continue with the single-life case now.

We next calculate the expectation and variance of Π_{t+h} conditional on the information available at time t , namely \mathcal{F}_t . First,

$$\begin{aligned} \mathbf{E}(\Pi_{t+h} | \mathcal{F}_t) = & \Pi - \mathbf{E}^{r, \lambda, t} \int_t^{t+h} \mathcal{D}^\mu P(r_s, \lambda_s, s) ds \\ & + \mathbf{E}^{r, \lambda, t} \int_t^{t+h} \pi_s (r_s F(r_s, s) + q(r_s, s) \sigma(r_s, s) F_r(r_s, s)) ds. \end{aligned} \quad (2.9)$$

Here, $\Pi_t = \Pi$ is known at time t , and $\mathbf{E}^{r, \lambda, t}$ denotes the conditional expectation given $r_t = r$ and $\lambda_t = \lambda$. Define the stochastic process Y_h for $h \geq 0$ by

$$Y_h = \Pi - \int_t^{t+h} \mathcal{D}^\mu P(r_s, \lambda_s, s) ds + \int_t^{t+h} \pi_s(r_s F(r_s, s) + q(r_s, s) \sigma(r_s, s) F_r(r_s, s)) ds. \quad (2.10)$$

Thus, $\mathbf{E}(\Pi_{t+h}|\mathcal{F}_t) = \mathbf{E}^{r,\lambda,t}Y_h$, and from (2.7), we have

$$\begin{aligned} \Pi_{t+h} = Y_h &+ \int_t^{t+h} \sigma(r_s, s)(\pi_s F_r(r_s, s) - P_r(r_s, \lambda_s, s)) dW_s \\ &- \int_t^{t+h} b(s)(\lambda_s - \underline{\lambda}) P_\lambda(r_s, \lambda_s, s) dW_s^\lambda + \int_t^{t+h} P(r_s, \lambda_s, s)(dN_s - \lambda_s ds). \end{aligned} \quad (2.11)$$

It follows that

$$\begin{aligned} \mathbf{Var}(\Pi_{t+h}|\mathcal{F}_t) &= \mathbf{E}((\Pi_{t+h} - \mathbf{E}Y_h)^2|\mathcal{F}_t) \\ &= \mathbf{E}^{r,\lambda,t}(Y_h - \mathbf{E}Y_h)^2 + \mathbf{E}^{r,\lambda,t} \int_t^{t+h} \sigma^2(r_s, s)(\pi_s F_r(r_s, s) - P_r(r_s, \lambda_s, s))^2 ds \\ &\quad + \mathbf{E}^{r,\lambda,t} \int_t^{t+h} b^2(s)(\lambda_s - \underline{\lambda})^2 P_\lambda^2(r_s, \lambda_s, s) ds + \mathbf{E}^{r,\lambda,t} \int_t^{t+h} \lambda_s P^2(r_s, \lambda_s, s) ds. \end{aligned} \quad (2.12)$$

We choose π_t in order to minimize the local variance $\lim_{h \rightarrow 0} \frac{1}{h} \mathbf{Var}(\Pi_{t+h}|\mathcal{F}_t)$, a dynamic measure of risk of the portfolio; therefore, $\pi_t = P_r(r_t, \lambda_t, t)/F_r(r_t, t)$. Under this assignment, the drift and local variance become, respectively,

$$\lim_{h \rightarrow 0} \frac{1}{h} (\mathbf{E}(\Pi_{t+h}|\mathcal{F}_t) - \Pi) = -\mathcal{D}^{\mu^Q} P(r, \lambda, t) + r P_r(r, \lambda, t) \frac{F(r, t)}{F_r(r, t)}, \quad (2.13)$$

and

$$\lim_{h \rightarrow 0} \frac{1}{h} \mathbf{Var}(\Pi_{t+h}|\mathcal{F}_t) = b^2(t)(\lambda - \underline{\lambda})^2 P_\lambda^2(r, \lambda, t) + \lambda P^2(r, \lambda, t). \quad (2.14)$$

Now, we come to pricing via the instantaneous Sharpe ratio. Because the minimum local variance in (2.14) is positive, the insurer is unable to completely hedge the risk of the pure endowment contract. Therefore, the price should reimburse the insurer for its risk, say, by a constant multiple α of the local standard deviation of the portfolio. It is this α that is the instantaneous Sharpe ratio.

From (2.14), we learn that the local standard deviation of the portfolio equals

$$\lim_{h \rightarrow 0} \sqrt{\frac{1}{h} \mathbf{Var}(\Pi_{t+h}|\mathcal{F}_t)} = \sqrt{b^2(t)(\lambda - \underline{\lambda})^2 P_\lambda^2(r, \lambda, t) + \lambda P^2(r, \lambda, t)}. \quad (2.15)$$

To determine the value (price) P , we set the drift of the portfolio equal to the short rate times the portfolio *plus* α times the local standard deviation. Thus, from (2.13) and (2.15), we have that P solves the equation

$$-\mathcal{D}^{\mu^Q} P + rP_r \frac{F}{F_r} = r\Pi + \alpha \sqrt{b^2(t)(\lambda - \underline{\lambda})^2 P_\lambda^2 + \lambda P^2}, \quad (2.16)$$

for some $0 \leq \alpha \leq \sqrt{\underline{\lambda}}$. Recall that $\Pi = -P + \pi F = -P + P_r F / F_r$. It follows that $P = P(r, \lambda, t)$ solves the non-linear pde given by

$$\begin{cases} P_t + \mu^Q P_r + \frac{1}{2} \sigma^2 P_{rr} + a P_\lambda + \frac{1}{2} b^2 (\lambda - \underline{\lambda})^2 P_{\lambda\lambda} - (r + \lambda) P \\ \quad = -\alpha \sqrt{b^2 (\lambda - \underline{\lambda})^2 P_\lambda^2 + \lambda P^2} \\ P(r, \lambda, T) = 1. \end{cases} \quad (2.17)$$

If we had been able to choose the investment strategy π so that the local standard deviation in (2.15) were identically zero (that is, if the risk were hedgeable), then the right-hand side of the pde in (2.17) would be zero, and we would have a linear differential equation of the Black-Scholes type. One can think of the right-hand side as adding a margin to the return of the portfolio because the pure endowment risk is not completely hedgeable due to the mortality risk. In addition to the unhedgeable mortality risk, the insurer also faces the somewhat diversifiable finite portfolio risk, that is, the risk that even if the hazard rate is deterministic, the actual number who survive until time T is different from expected. This risk is clearly present when selling a pure endowment to a single individual, but it is also present to some extent in any portfolio of finite size. In Section 4.4, we decompose the risk loading in the price due to the finite portfolio risk and due to the stochastic mortality risk.

If there were no risk loading, that is, if $\alpha = 0$, then the price is such that the expected return on the price is $r + \lambda$. The rate r arises from the riskless money market, and λ arises from the expected release of reserves as individuals die. If $\alpha > 0$, then the expected return on the price is greater than $r + \lambda$. Therefore, α , the Sharpe ratio, measures the degree to which the insurer's total expected return is in excess of $r + \lambda$, as a proportion of the standard deviation of the return.

Before moving on to the following sections where we study properties of the solution of (2.17), we show that we can simplify P greatly. Indeed, $P(r, \lambda, t) = F(r, t)\varphi(\lambda, t)$, in which F is the price of the T -bond and solves (2.5), and φ solves the non-linear pde

$$\begin{cases} \varphi_t + a\varphi_\lambda + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\varphi_{\lambda\lambda} - \lambda\varphi = -\alpha\sqrt{b^2(\lambda - \underline{\lambda})^2\varphi_\lambda^2 + \lambda\varphi^2}, \\ \varphi(\lambda, T) = 1. \end{cases} \quad (2.18)$$

The existence of a solution to (2.18) follows from standard techniques; see, for example, Walter (1970, Chapter IV, Section 36). A comparison result (see Section 3 of this paper) demonstrates that the solution is unique.

The factorization $P = F\varphi$ is reminiscent of the standard actuarial method of pricing pure endowments in that φ represents the probability of paying the mortality-contingent claim, that is, the probability that the individual survives. We will see in Section 3.1 that we can interpret φ as a risk-adjusted survival probability. As mentioned at the end of Section 2.1, this factorization arises under other commonly used models for bond prices, as long as the risk driving the bond price (W in our case) is independent of the risk driving the stochastic hazard rate, W^λ .

Consider the special case for which $b \equiv 0$, that is, λ is deterministic. Suppose $\lambda(t)$ is the solution of $d\lambda = a(\lambda, s)ds$ with initial value $\lambda_0 = \lambda$; then, (2.18) becomes the linear ordinary differential equation

$$\varphi'(t) - (\lambda(t) - \alpha\sqrt{\lambda(t)})\varphi(t) = 0, \quad \varphi(T) = 1, \quad (2.19)$$

whose solution is

$$\varphi(t) = e^{-\int_t^T (\lambda(s) - \alpha\sqrt{\lambda(s)})ds}, \quad (2.20)$$

a type of probability of survival because we can think of $\lambda(t) - \alpha\sqrt{\lambda(t)} > 0$ as a modified hazard rate. Indeed, note that if $\alpha = 0$, then (2.20) is the physical probability that a person alive at time t survives to time T , and as α increases, φ increases. Therefore, we interpret (2.20) as a risk-adjusted probability of survival, in which α controls the degree to which we adjust (that is, increase) the physical probability of survival. When $b \neq 0$, a similar phenomenon occurs (see Theorem 3.8 below), and for this reason, we refer to the solution P of (2.17) as the risk-adjusted price for the pure endowment.

Recall that we assume that $0 \leq \alpha \leq \sqrt{\underline{\lambda}}$. For the case of deterministic hazard, this implies that $\lambda(t) - \alpha\sqrt{\lambda(t)} > 0$ for all $t \in [0, T]$, so that $0 \leq \varphi(t) \leq 1$ in (2.20). For stochastic hazard, we show in the next section that $0 \leq \varphi(\lambda, t) \leq 1$ for all $(\lambda, t) \in (\underline{\lambda}, \infty) \times [0, T]$. It follows that, in general, F is an upper bound for the price. Observe that F is a natural upper bound for the price because it is the price we would charge if we knew the person could not die before T .

3. Qualitative Properties of the Risk-Adjusted Price

In this section, we discuss qualitative properties of the risk-adjusted price P in (2.17) and show that it shares many properties with the (static) standard deviation premium principle (Gerber, 1979). To begin, we have the following proposition.

Proposition 3.1. *Suppose P^c is the price, as determined by the method in Section 2, for a pure endowment with payment $c \geq 0$ at time T if the individual is alive. Then, $P^c = cP$, in which P is the risk-adjusted price for a payment of 1 at time T if the individual is alive.*

Proof. In the derivation of P for (2.17), it is clear that if we derive the price P^c , then (2.17) still applies for determining P^c with the terminal condition $P^c(r, \lambda, T) = c$. We can, then, write $P^c = cF\varphi = cP$. \square

Proposition 3.1 parallels the following well known fact concerning the (static) standard deviation premium principle H : If we define the standard deviation premium principle, as applied to a random variable X , by

$$H(X) = \mathbf{E}X + \alpha\sqrt{\mathbf{Var}X}, \quad (3.1)$$

then $H(cX) = cH(X)$ for $c \geq 0$.

In what follows, we show that $0 \leq P \leq F$ and $P_\lambda \leq 0$, and we examine how the price P responds to changes in the model parameters. To this end, we need a comparison principle (Walter, 1970, Section 28). We begin by stating a relevant one-sided Lipschitz condition along with growth conditions. We require that the function $g = g(\lambda, t, v, p)$ satisfies the following one-sided Lipschitz condition: For $v > w$,

$$g(\lambda, t, v, p) - g(\lambda, t, w, q) \leq c(\lambda, t)(v - w) + d(\lambda, t)|p - q|, \quad (3.2)$$

with growth conditions on c and d given by

$$0 \leq c(\lambda, t) \leq K(1 + (\ln(\lambda - \underline{\lambda}))^2), \text{ and } 0 \leq d(\lambda, t) \leq K(\lambda - \underline{\lambda})(1 + |\ln(\lambda - \underline{\lambda})|), \quad (3.3)$$

for some constant $K \geq 0$, and for all $(\lambda, t) \in (\underline{\lambda}, \infty) \times [0, T]$. Throughout this paper, we rely on the following useful comparison principle, which we obtain from Walter (1970, Section 28).

Theorem 3.2. *Let $G = (\underline{\lambda}, \infty) \times [0, T]$, and denote by \mathcal{G} the collection of functions on G that are twice-differentiable in their first variable and once-differentiable in their second.*

Define a differential operator \mathcal{L} on \mathcal{G} by

$$\mathcal{L}v = v_t + \frac{1}{2}b^2(t)(\lambda - \underline{\lambda})^2 v_{\lambda\lambda} + g(\lambda, t, v, v_\lambda), \quad (3.4)$$

in which g satisfies (3.2) and (3.3). Suppose $v, w \in \mathcal{G}$ are such that there exists a constant $K \geq 0$ with $v \leq e^{K(\ln(\lambda - \underline{\lambda}))^2}$ and $w \geq -e^{K(\ln(\lambda - \underline{\lambda}))^2}$ for large λ and for λ close to $\underline{\lambda}$. Then, if (a) $\mathcal{L}v \geq \mathcal{L}w$ on G , and if (b) $v(\lambda, T) \leq w(\lambda, T)$ for all $\lambda > \underline{\lambda}$, then $v \leq w$ on G .

Proof. Transform the variables λ and t in (3.4) to $y = \ln(\lambda - \underline{\lambda})$ and $\tau = T - t$, and write $\tilde{v}(y, \tau) = v(\lambda, t)$, etc. Under this transformation, (3.4) becomes

$$\mathcal{L}\tilde{v} = -\tilde{v}_\tau + \frac{1}{2}\tilde{b}^2(\tau)\tilde{v}_{yy} + \tilde{h}(y, \tau, \tilde{v}, \tilde{v}_y), \quad (3.5)$$

in which $\tilde{h}(y, \tau, \tilde{v}, \tilde{p}) = -\frac{1}{2}\tilde{b}^2(\tau)\tilde{p} + \tilde{g}(y, \tau, \tilde{v}, \tilde{p})$, and \tilde{v} is a differential function on $\mathbf{R} \times [0, T]$. Note that $\varphi_\lambda = e^{-y}\tilde{\varphi}_y$, so $p = e^{-y}\tilde{p}$ in going from g to \tilde{g} . The differential operator in (3.5) is of the form considered by Walter (1970, Section 28, pages 213-215); see that reference for the proof of our assertion.

The remaining item to consider is the form of the growth conditions in the original variables λ and t . From Walter (1970), we know that analog of (3.2) and (3.3) for \tilde{h} are

$$\tilde{h}(y, \tau, \tilde{v}, \tilde{p}) - \tilde{h}(y, \tau, \tilde{w}, \tilde{q}) \leq \tilde{c}(y, \tau)(\tilde{v} - \tilde{w}) + \tilde{d}(y, \tau)|\tilde{p} - \tilde{q}|, \quad (3.6)$$

with

$$0 \leq \tilde{c}(y, \tau) \leq K(1 + y^2), \text{ and } 0 \leq \tilde{d}(y, \tau) \leq K(1 + |y|). \quad (3.7)$$

Under the original variables, the right-hand side of (3.6) becomes $c(\lambda, t)(v - w) + d(\lambda, t)|p - q|$, in which $c(\lambda, t) = \tilde{c}(y, \tau)$ and $d(\lambda, t) = \tilde{d}(y, \tau)e^y$ because $\tilde{p} = e^y p$. Therefore, $\tilde{d}(y, \tau) \leq K(1 + |y|)$ becomes $d(\lambda, t) \leq Ke^y(1 + |y|) = K(\lambda - \underline{\lambda})(1 + |\ln(\lambda - \underline{\lambda})|)$. \square

As a lemma for results to follow, we show that the differential operator associated with our problem satisfies the hypotheses of Theorem 3.2.

Lemma 3.3. *If we define g by*

$$g(\lambda, t, v, p) = a(\lambda, t)p - \lambda v + \alpha\sqrt{b^2(t)(\lambda - \underline{\lambda})^2 p^2 + \lambda v^2}, \quad (3.8)$$

then g satisfies the one-sided Lipschitz condition (3.2) on G . Furthermore, if $|a(\lambda, t)| \leq K(\lambda - \underline{\lambda})(1 + |\ln(\lambda - \underline{\lambda})|)$, then (3.3) holds.

Proof. Suppose $v > w$, then

$$\begin{aligned}
g(\lambda, t, v, p) - g(\lambda, t, w, q) &= a(\lambda, t)(p - q) - \lambda(v - w) \\
&\quad + \alpha \left\{ \sqrt{b^2(t)(\lambda - \underline{\lambda})^2 p^2 + \lambda v^2} - \sqrt{b^2(t)(\lambda - \underline{\lambda})^2 q^2 + \lambda w^2} \right\} \\
&\leq (|a(\lambda, t)| + \alpha b(t)(\lambda - \underline{\lambda}))|p - q| - \left(\lambda - \alpha \sqrt{\lambda} \right) (v - w) \\
&\leq (|a(\lambda, t)| + \alpha b(t)(\lambda - \underline{\lambda}))|p - q|.
\end{aligned} \tag{3.9}$$

Recall that $\alpha \leq \sqrt{\underline{\lambda}}$. Also, we use the fact that if $A \geq B$, then $\sqrt{C^2 + A^2} - \sqrt{C^2 + B^2} \leq A - B$, as we demonstrate below in Lemma 4.5. Thus, (3.2) holds with $c(\lambda, t) = 0$ and $d(\lambda, t) = |a(\lambda, t)| + \alpha b(t)(\lambda - \underline{\lambda})$. Note that d satisfies (3.3) if $|a(\lambda, t)| \leq K(\lambda - \underline{\lambda})(1 + |\ln(\lambda - \underline{\lambda})|)$. \square

Assumption 3.4. Henceforth, we assume that the drift a satisfies the growth condition in the hypothesis of Lemma 3.3. For later purposes (for example, see Theorem 3.7), we also assume that a_λ is Hölder continuous and satisfies the growth condition $|a_\lambda| \leq K(1 + (\ln(\lambda - \underline{\lambda}))^2)$.

In the next two subsections, we apply Theorem 3.2 and Lemma 3.3 repeatedly to determine qualitative properties of the risk-adjusted premium P .

3.1. Interpreting φ as a Survival Probability

In our first applications of Theorem 3.2, we show that φ shares two important properties with physical survival probabilities, namely that $0 \leq \varphi \leq 1$ and $\varphi_\lambda \leq 0$. In fact, we have an even tighter upper bound on φ by observing that $\underline{\lambda}$ is a lower bound for λ .

Theorem 3.5. $0 \leq \varphi(\lambda, t) \leq e^{-(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}})(T-t)}$ for $(\lambda, t) \in G$.

Proof. Define the differential operator \mathcal{L} on \mathcal{G} by (3.4) with g given in (3.8). Because φ solves (2.18), we have $\mathcal{L}\varphi = 0$. Also,

$$\begin{aligned}
\mathcal{L}e^{-(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}})(T-t)} &= \left(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}} \right) e^{-(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}})(T-t)} - \left(\lambda - \alpha \sqrt{\lambda} \right) e^{-(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}})(T-t)} \\
&\propto \left(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}} \right) - \left(\lambda - \alpha \sqrt{\lambda} \right) \leq 0.
\end{aligned} \tag{3.10}$$

Because $\mathcal{L}e^{-(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}})(T-t)} \leq \mathcal{L}\varphi$ and $\varphi(\lambda, T) = 1 = e^{-(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}})(T-T)}$, Theorem 3.2 and Lemma 3.3 imply that $\varphi \leq e^{-(\underline{\lambda} - \alpha \sqrt{\underline{\lambda}})(T-t)}$ on G .

Similarly, denote by $\mathbf{0}$ the function that is identically 0 on G ; then, $\mathcal{L}\mathbf{0} = 0 = \mathcal{L}\varphi$. Because additionally $\varphi(\lambda, T) = 1$, Theorem 3.2 and Lemma 3.3 imply that $0 \leq \varphi$ on G . \square

The upper bound in Theorem 3.5 is tight. Indeed, suppose $\lambda \equiv \underline{\lambda} + \epsilon$ for some constant $\epsilon > 0$. Then, the solution to (2.18) is given by $e^{-((\lambda + \epsilon) - \alpha\sqrt{\lambda + \epsilon})(T-t)}$, which can be made arbitrarily close to $e^{-(\lambda - \alpha\sqrt{\underline{\lambda}})(T-t)}$ (uniformly on $[0, T]$) by choosing ϵ small enough.

We have the following corollary of Theorem 3.5 that gives us natural bounds for the risk-adjusted price. F is a natural bound for the price because it is the price of a default-free bond, that is, a bond that pays regardless of whether the individual is alive.

Corollary 3.6. $0 \leq P(r, \lambda, t) \leq F(r, t)$ for $(r, \lambda, t) \in \mathbf{R}^+ \times G$.

Proof. Because $P(r, \lambda, t) = F(r, t)\varphi(\lambda, t)$ and $e^{-(\lambda - \alpha\sqrt{\underline{\lambda}})(T-t)} \leq 1$, the result is immediate from Theorem 3.5. \square

We end this subsection with a proof that $\varphi_\lambda \leq 0$. This result is intuitive for physical survival probabilities in that if the current hazard rate λ increases, then the probability of surviving until time T decreases.

Theorem 3.7. $\varphi_\lambda(\lambda, t) \leq 0$ for $(\lambda, t) \in G$.

Proof. To prove this assertion, we apply a modified version of Theorem 3.2 to the special case of comparing φ_λ with the zero function $\mathbf{0}$. From Walter (1970, Section 28, pages 213-215), we see that we only need to verify that (3.2) holds for $v > 0 = w = q$. First, differentiate φ 's equation with respect to λ to get an equation for $f = \varphi_\lambda$.

$$\begin{cases} f_t + (a_\lambda - \lambda)f + (a + b^2(\lambda - \underline{\lambda}))f_\lambda + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2 f_{\lambda\lambda} - \varphi \\ \quad = -\alpha \frac{b^2(\lambda - \underline{\lambda})f^2 + b^2(\lambda - \underline{\lambda})^2 f f_\lambda + \frac{1}{2}\varphi^2 + \lambda\varphi f}{\sqrt{b^2(\lambda - \underline{\lambda})^2 f^2 + \lambda\varphi^2}}, \\ f(\lambda, T) = 0. \end{cases} \quad (3.11)$$

Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with g given by

$$g(\lambda, t, v, p) = (a_\lambda - \lambda)v + (a + b^2(\lambda - \underline{\lambda}))p - \varphi + \alpha \frac{b^2(\lambda - \underline{\lambda})v^2 + b^2(\lambda - \underline{\lambda})^2 vp + \frac{1}{2}\varphi^2 + \lambda\varphi v}{\sqrt{b^2(\lambda - \underline{\lambda})^2 v^2 + \lambda\varphi^2}}. \quad (3.12)$$

To apply a modified version of Theorem 3.2, verify that (3.2) and (3.3) hold for $v > 0 = w = q$. It is not difficult to show that in this case,

$$\begin{aligned} g(\lambda, t, v, p) - g(\lambda, t, 0, 0) &\leq (|a_\lambda| + \alpha b - (\lambda - \alpha\sqrt{\underline{\lambda}}))v + (|a| + b^2(\lambda - \underline{\lambda}) + \alpha b(\lambda - \underline{\lambda}))|p| \\ &\leq (|a_\lambda| + \alpha b)v + (|a| + b^2(\lambda - \underline{\lambda}) + \alpha b(\lambda - \underline{\lambda}))|p|. \end{aligned} \quad (3.13)$$

Thus, by Assumption 3.4, g satisfies (3.2) with the corresponding c and d satisfying the growth conditions in (3.3).

Next, note that because $f = \varphi_\lambda$ satisfies (3.11), $\mathcal{L}f = 0$. Also, $\mathcal{L}\mathbf{0} = \varphi(-1 + \alpha/(2\sqrt{\lambda})) \leq 0$ because $\lambda \geq \underline{\lambda} \geq \alpha^2$. These observations, together with $f(\lambda, T) = 0$, imply that $f = \varphi_\lambda \leq 0$ on G . \square

As in the relationship between Theorem 3.5 and Corollary 3.6, we have the immediate corollary of Theorem 3.7 that $P_\lambda \leq 0$ on G .

3.2. Comparative Statics for P

Our next results show that as we vary the model parameters, the price P responds consistently with what we expect.

Theorem 3.8. *Suppose $0 \leq \alpha_1 < \alpha_2 \leq \sqrt{\underline{\lambda}}$, and let P^{α_i} be the solution to (2.17) with $\alpha = \alpha_i$, for $i = 1, 2$. Then, $P^{\alpha_1}(r, \lambda, t) \leq P^{\alpha_2}(r, \lambda, t)$ for all $(r, \lambda, t) \in \mathbf{R}^+ \times G$.*

Proof. Because F in (2.5) is independent of α , it is enough to show that $\varphi^{\alpha_1} \leq \varphi^{\alpha_2}$ on G , in which φ^{α_i} has the obvious meaning. Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) and (3.8) with $\alpha = \alpha_1$. Because φ^{α_1} solves (2.18) with $\alpha = \alpha_1$, we have $\mathcal{L}\varphi^{\alpha_1} = 0$. Also,

$$\begin{aligned} \mathcal{L}\varphi^{\alpha_2} &= \varphi_t^{\alpha_2} + a\varphi_\lambda^{\alpha_2} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\varphi_{\lambda\lambda}^{\alpha_2} - \lambda\varphi^{\alpha_2} + \alpha_1\sqrt{b^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{\alpha_2})^2 + \lambda(\varphi^{\alpha_2})^2} \\ &= -(\alpha_2 - \alpha_1)\sqrt{b^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{\alpha_2})^2 + \lambda(\varphi^{\alpha_2})^2} \leq 0 = \mathcal{L}\varphi^{\alpha_1}. \end{aligned} \tag{3.14}$$

In addition, both φ^{α_1} and φ^{α_2} satisfy the terminal condition $\varphi^{\alpha_i}(\lambda, T) = 1$. Thus, Theorem 3.2 and Lemma 3.3 imply that $\varphi^{\alpha_1} \leq \varphi^{\alpha_2}$ on G . \square

Theorem 3.8 states that as the parameter α increases, the price P^α increases; this result justifies the use of the phrase *risk parameter* when referring to α . It is clear that the price obtained from the standard deviation premium principle, as defined in (3.1), also increases with α . We have the following corollary to Theorem 3.8.

Corollary 3.9. *Let P^{α_0} be the solution to (2.17) with $\alpha = 0$; then, $P^{\alpha_0} \leq P^\alpha$ for all $0 \leq \alpha \leq \sqrt{\underline{\lambda}}$, and we can express the lower bound P^{α_0} as follows: $P^{\alpha_0}(r, \lambda, t) = F(r, t)\varphi^{\alpha_0}(\lambda, t)$, in which φ^{α_0} is given by the physical probability of survival, namely*

$$\varphi^{\alpha_0}(\lambda, t) = \mathbf{E} \left[e^{-\int_t^T \lambda_s ds} \middle| \lambda_t = \lambda \right], \tag{3.15}$$

where λ_s follows the process given in (2.1).

Proof. Theorem 3.8 implies that $P^{\alpha 0} \leq P^\alpha$ for all $0 \leq \alpha \leq \sqrt{\underline{\lambda}}$, and by substituting $\alpha = 0$ in (2.18), the Feynman-Kac Theorem (Karatzas and Shreve, 1991) implies the representation of $\varphi^{\alpha 0}$ in (3.15). \square

Corollary 3.9 justifies the use of the phrase *risk-adjusted* price when referring to P^α because $P^\alpha \geq P^{\alpha 0}$, and we can interpret $P^{\alpha 0}$ as a *risk-neutral* price due to the fact that $P^{\alpha 0}$ equals the product of the bond price and the probability of paying the pure endowment. We call $P^\alpha - P^{\alpha 0}$ the *risk charge* that compensates the insurer for both the finite portfolio risk and the stochastic mortality risk. In Section 4.4, we decompose the risk charge into these two components after we study the price for a portfolio of n pure endowment risks.

Next, we examine how the risk-adjusted price P varies with the drift and volatility of the stochastic hazard rate.

Theorem 3.10. *Suppose $a_1(\lambda, t) \leq a_2(\lambda, t)$ on G , and let P^{a_i} denote the solution to (2.17) with $a = a_i$, for $i = 1, 2$. Then, $P^{a_1}(r, \lambda, t) \geq P^{a_2}(r, \lambda, t)$ for all $(r, \lambda, t) \in \mathbf{R}^+ \times G$.*

Proof. Because F in (2.5) is independent of a , it is enough to show that $\varphi^{a_1} \geq \varphi^{a_2}$, in which φ^{a_i} has the obvious meaning. Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) and (3.8) with $a = a_1$. Because φ^{a_1} solves (2.18) with $a = a_1$, we have $\mathcal{L}\varphi^{a_1} = 0$. Also, because $\varphi_\lambda^{a_2} \leq 0$, then

$$\begin{aligned} \mathcal{L}\varphi^{a_2} &= \varphi_t^{a_2} + a_1\varphi_\lambda^{a_2} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\varphi_{\lambda\lambda}^{a_2} - \lambda\varphi^{a_2} + \alpha\sqrt{b^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{a_2})^2 + \lambda(\varphi^{a_2})^2} \\ &= -(a_2 - a_1)\varphi_\lambda^{a_2} \geq 0 = \mathcal{L}\varphi^{a_1}. \end{aligned} \quad (3.16)$$

In addition, both φ^{a_1} and φ^{a_2} satisfy the terminal condition $\varphi^{a_i}(\lambda, T) = 1$. Theorem 3.2 and Lemma 3.3 imply that $\varphi^{a_1} \geq \varphi^{a_2}$ on G . \square

The result parallel to Theorem 3.10 for the standard deviation premium principle in (3.1) is that H decreases as $\mathbf{E}X$ decreases.

Theorem 3.11. *Suppose $0 \leq b_1(t) \leq b_2(t)$ on $[0, T]$, and let P^{b_i} denote the solution to (2.17) with $b = b_i$, for $i = 1, 2$. If $P_{\lambda\lambda}^{b_i} \geq 0$ for $i = 1$ or 2 , then $P^{b_1}(r, \lambda, t) \leq P^{b_2}(r, \lambda, t)$ for all $(r, \lambda, t) \in \mathbf{R}^+ \times G$.*

Proof. Because F in (2.5) is independent of b , it is enough to show that $\varphi^{b_1} \leq \varphi^{b_2}$ if $\varphi_{\lambda\lambda}^{b_i} \geq 0$, in which φ^{b_i} has the obvious meaning. Suppose $\varphi_{\lambda\lambda}^{b_2} \geq 0$, and define a differential operator \mathcal{L} on \mathcal{G} by (3.4) and (3.8) with $b = b_1$. Because φ^{b_1} solves (2.18) with $b = b_1$, we have $\mathcal{L}\varphi^{b_1} = 0$. Also,

$$\begin{aligned}
\mathcal{L}\varphi^{b_2} &= \varphi_t^{b_2} + a\varphi_\lambda^{b_2} + \frac{1}{2}b_1^2(\lambda - \underline{\lambda})^2\varphi_{\lambda\lambda}^{b_2} - \lambda\varphi^{b_2} + \alpha\sqrt{b_1^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{b_2})^2 + \lambda(\varphi^{b_2})^2} \\
&= -\frac{1}{2}(b_2^2 - b_1^2)(\lambda - \underline{\lambda})^2\varphi_{\lambda\lambda}^{b_2} \\
&\quad - \alpha\left\{\sqrt{b_2^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{b_2})^2 + \lambda(\varphi^{b_2})^2} - \sqrt{b_1^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{b_2})^2 + \lambda(\varphi^{b_2})^2}\right\} \\
&\leq 0 = \mathcal{L}\varphi^{b_1}.
\end{aligned} \tag{3.17}$$

In addition, both φ^{b_1} and φ^{b_2} satisfy the terminal condition $\varphi^{b_i}(\lambda, T) = 1$. Thus, Theorem 3.2 and Lemma 3.3 imply that $\varphi^{b_1} \leq \varphi^{b_2}$ on G . In a similar fashion, we obtain the same result if $\varphi_{\lambda\lambda}^{b_1} \geq 0$. \square

From Theorem 3.11, we see that if φ is convex with respect to λ , then the risk-adjusted price increases if the volatility on the stochastic hazard rate increases. The parallel result for the standard deviation premium principle in (3.1) is that H increases as $\mathbf{Var}X$ increases.

4. The Risk-Adjusted Price for a Portfolio of Risks

In this section, we study properties of the price $P^{(n)} = P^{(n)}(r, \lambda, t)$ for n conditionally independent and identically distributed pure endowment risks. First, we present the equation that $P^{(n)}$ solves. In Sections 4.1 and 4.2, we parallel the results of Sections 3.1 and 3.2, respectively. In Section 4.3, we show that $P^{(n)}$ is subadditive. In Section 4.4, we show that the risk charge per person decreases as the n increases. We also show that if the hazard rate is deterministic, then the risk charge per person goes to zero as n goes to infinity. Moreover, we show that if the hazard rate is stochastic, then the risk charge per person is positive as n goes to infinity, which reflects the fact that the stochastic mortality risk is not diversifiable.

As discussed in the paragraph preceding equation (2.9), when an individual dies, the portfolio value Π jumps by $P^{(n)} - P^{(n-1)}$. By paralleling the derivation of (2.17), one can show that $P^{(n)}$ solves the non-linear pde given by

$$\begin{cases} P_t^{(n)} + \mu^Q P_r^{(n)} + \frac{1}{2}\sigma^2 P_{rr}^{(n)} + aP_\lambda^{(n)} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2 P_{\lambda\lambda}^{(n)} - rP^{(n)} - n\lambda(P^{(n)} - P^{(n-1)}) \\ \quad = -\alpha\sqrt{b^2(\lambda - \underline{\lambda})^2(P_\lambda^{(n)})^2 + n\lambda(P^{(n)} - P^{(n-1)})^2} \\ P^{(n)}(r, \lambda, T) = n. \end{cases} \tag{4.1}$$

The initial value in this recursion is $P^{(0)} \equiv 0$, and the price P as defined by (2.17) is $P^{(1)}$.

As in Section 2, we can multiplicatively separate the variables r and λ in $P^{(n)}$. Indeed, $P^{(n)}(r, \lambda, t) = F(r, t)\varphi^{(n)}(\lambda, t)$, in which F solves (2.5) and $\varphi^{(n)}$ solves the recursion

$$\begin{cases} \varphi_t^{(n)} + a\varphi_\lambda^{(n)} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\varphi_{\lambda\lambda}^{(n)} - n\lambda(\varphi^{(n)} - \varphi^{(n-1)}) \\ \quad = -\alpha\sqrt{b^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{(n)})^2 + n\lambda(\varphi^{(n)} - \varphi^{(n-1)})^2}, \\ \varphi^{(n)}(\lambda, T) = n, \end{cases} \quad (4.2)$$

with initial value $\varphi^{(0)} \equiv 0$. Note that φ in (2.18) equals $\varphi^{(1)}$.

Throughout this section, we apply Theorem 3.2 with $g = g_n$ defined by (4.3) below. We have the following lemma whose proof we omit because it parallels that of Lemma 3.3.

Lemma 4.1. *Define g_n , for $n \geq 1$, by*

$$g_n(\lambda, t, v, p) = a(\lambda, t)p - n\lambda(v - \varphi^{(n-1)}) + \alpha\sqrt{b^2(\lambda - \underline{\lambda})^2p^2 + n\lambda(v - \varphi^{(n-1)})^2}, \quad (4.3)$$

in which $\varphi^{(n-1)}$ solves (4.2) with n replaced by $n - 1$. Then, g_n satisfies the one-sided Lipschitz condition (3.2) on G . Furthermore, if $|a(\lambda, t)| \leq K(\lambda - \underline{\lambda})(1 + |\ln(\lambda - \underline{\lambda})|)$, then (3.3) holds.

4.1. Interpreting $\frac{1}{n}\varphi^{(n)}$ as a Survival Probability

In the first application of Theorem 3.2 and Lemma 4.1, we show that $0 \leq \frac{1}{n}\varphi^{(n)} \leq e^{-(\underline{\lambda} - \alpha\sqrt{\underline{\lambda}})(T-t)}$ for $n \geq 1$.

Theorem 4.2. $0 \leq \varphi^{(n)}(\lambda, t) \leq ne^{-(\underline{\lambda} - \alpha\sqrt{\underline{\lambda}})(T-t)}$ for $(\lambda, t) \in G$ and for $n \geq 0$.

Proof. We proceed by induction to prove the upper bound. For the ease of presentation in this proof, define $h(t) = e^{-(\underline{\lambda} - \alpha\sqrt{\underline{\lambda}})(T-t)}$. It is clear that the inequality holds for $n = 0$ because $\varphi^{(0)} \equiv 0$. For $n \geq 1$, assume that $0 \leq \varphi^{(n-1)} \leq (n-1)h(t)$, and show that $0 \leq \varphi^{(n)} \leq nh(t)$.

Define the differential operator \mathcal{L} on \mathcal{G} by (3.4) with $g = g_n$ from (4.3). Because $\varphi^{(n)}$ solves (4.2), we have $\mathcal{L}\varphi^{(n)} = 0$. Also,

$$\begin{aligned} \mathcal{L}nh(t) &= n(\underline{\lambda} - \alpha\sqrt{\underline{\lambda}})h(t) - (n\lambda - \alpha\sqrt{n\lambda})(nh(t) - \varphi^{(n-1)}) \\ &\leq n(\underline{\lambda} - \alpha\sqrt{\underline{\lambda}})h(t) - (n\lambda - \alpha\sqrt{n\lambda})(n - (n-1))h(t) \\ &= [n(\underline{\lambda} - \alpha\sqrt{\underline{\lambda}}) - (n\lambda - \alpha\sqrt{n\lambda})]h(t) \leq 0. \end{aligned} \quad (4.4)$$

Because $\mathcal{L}ne^{-(\underline{\lambda}-\alpha\sqrt{\underline{\lambda}})(T-t)} \leq \mathcal{L}\varphi^{(n)}$ and $\varphi^{(n)}(\lambda, T) = n$, Theorem 3.2 and Lemma 4.1 imply that $\varphi^{(n)} \leq ne^{-(\underline{\lambda}-\alpha\sqrt{\underline{\lambda}})(T-t)}$ on G .

Similarly, we prove the lower bound via induction. Suppose that $\varphi^{(n-1)} \geq 0$ for $n \geq 1$, and show that $\varphi^{(n)} \geq 0$. Let \mathcal{L} be the differential operator from the first part of this proof, and denote by $\mathbf{0}$ the function that is identically 0 on G ; then, $\mathcal{L}\mathbf{0} = (n\lambda + \alpha\sqrt{n\lambda})\varphi^{(n-1)} \geq 0$. Because $\mathcal{L}\mathbf{0} \geq \mathcal{L}\varphi^{(n)}$ and $\varphi^{(n)}(\lambda, T) = n$, Theorem 3.2 and Lemma 4.1 imply that $0 \leq \varphi^{(n)}$ on G . \square

An immediate corollary of Theorem 4.2 is that $0 \leq P^{(n)} \leq nF$, as in Corollary 3.6, so, once again the price per risk $\frac{1}{n}P^{(n)}$ lies between 0 and F . We end this subsection by extending Theorem 3.7, namely $\varphi_\lambda \leq 0$, to an arbitrary number of risks. We first present a lemma which we use in its proof.

Lemma 4.3. $\varphi^{(n)}(\lambda, t) \geq \varphi^{(n-1)}(\lambda, t)$ for $(\lambda, t) \in G$ and for $n \geq 1$.

Proof. We proceed by induction. This inequality is true for $n = 1$ because $\varphi^{(1)} = \varphi \geq 0 = \varphi^{(0)}$ by Theorem 3.5. For $n \geq 2$, assume that $\varphi^{(n-1)} \geq \varphi^{(n-2)}$, and show that $\varphi^{(n)} \geq \varphi^{(n-1)}$.

Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with $g = g_n$ from (4.3). Because $\varphi^{(n)}$ solves (4.2), we have $\mathcal{L}\varphi^{(n)} = 0$. Also,

$$\begin{aligned} \mathcal{L}\varphi^{(n-1)} &= (n-1)\lambda \left(\varphi^{(n-1)} - \varphi^{(n-2)} \right) - n\lambda \left(\varphi^{(n-1)} - \varphi^{(n-1)} \right) \\ &\quad + \alpha \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\varphi_\lambda^{(n-1)} \right)^2 + n\lambda \left(\varphi^{(n-1)} - \varphi^{(n-1)} \right)^2} \\ &\quad - \alpha \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\varphi_\lambda^{(n-1)} \right)^2 + (n-1)\lambda \left(\varphi^{(n-1)} - \varphi^{(n-2)} \right)^2} \\ &\geq (n-1)\lambda \left(\varphi^{(n-1)} - \varphi^{(n-2)} \right) + \alpha b(\lambda - \underline{\lambda}) \left| \varphi_\lambda^{(n-1)} \right| \\ &\quad - \alpha \left(b(\lambda - \underline{\lambda}) \left| \varphi_\lambda^{(n-1)} \right| + \sqrt{(n-1)\lambda} \left(\varphi^{(n-1)} - \varphi^{(n-2)} \right) \right) \\ &= \left((n-1)\lambda - \alpha\sqrt{(n-1)\lambda} \right) \left(\varphi^{(n-1)} - \varphi^{(n-2)} \right) \geq 0 = \mathcal{L}\varphi^{(n)}. \end{aligned} \tag{4.5}$$

Note that the first inequality follows from the subadditivity of the square root, that is, $\sqrt{A^2 + B^2} \leq |A| + |B|$. We also use the induction hypothesis that $\varphi^{(n-1)} \geq \varphi^{(n-2)}$. In addition, $\varphi^{(n)}(\lambda, T) = n > n-1 = \varphi^{(n-1)}(\lambda, T)$. Thus, Theorem 3.2 and Lemma 4.1 imply that $\varphi^{(n)} \geq \varphi^{(n-1)}$ on G . \square

Lemma 4.3 is interesting in its own right because it confirms our intuition that $P^{(n)}$ increases with the number of policyholders n .

Theorem 4.4. $\varphi_\lambda^{(n)}(\lambda, t) \leq 0$ for $(\lambda, t) \in G$ and for $n \geq 0$.

Proof. We proceed by induction. We know that the inequality holds when $n = 0$ and 1 . For $n \geq 2$, assume that $\varphi_\lambda^{(n-1)} \leq 0$, and show that $\varphi_\lambda^{(n)} \leq 0$. As in the proof of Theorem 3.7, to prove this assertion, we apply a modified version of Theorem 3.2 to the special case of comparing $\varphi_\lambda^{(n)}$ with the zero function $\mathbf{0}$. First, differentiate $\varphi^{(n)}$'s equation with respect to λ to get an equation for $f^{(n)} = \varphi_\lambda^{(n)}$.

$$\left\{ \begin{array}{l} f_t^{(n)} + (a_\lambda - n\lambda)f^{(n)} + n\lambda f^{(n-1)} + (a + b^2(\lambda - \underline{\lambda}))f_\lambda^{(n)} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2 f_{\lambda\lambda}^{(n)} \\ \quad - n(\varphi^{(n)} - \varphi^{(n-1)}) \\ = -\alpha \frac{b^2(\lambda - \underline{\lambda})(f^{(n)})^2 + b^2(\lambda - \underline{\lambda})^2 f^{(n)} f_\lambda^{(n)} + \frac{1}{2}n(\varphi^{(n)} - \varphi^{(n-1)})^2}{\sqrt{b^2(\lambda - \underline{\lambda})^2 (f^{(n)})^2 + n\lambda(\varphi^{(n)} - \varphi^{(n-1)})^2}} \\ \quad - \alpha \frac{n\lambda(\varphi^{(n)} - \varphi^{(n-1)})(f^{(n)} - f^{(n-1)})}{\sqrt{b^2(\lambda - \underline{\lambda})^2 (f^{(n)})^2 + n\lambda(\varphi^{(n)} - \varphi^{(n-1)})^2}}, \\ f^{(n)}(\lambda, T) = 0. \end{array} \right. \quad (4.6)$$

Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with $g = g_n$ given by

$$\begin{aligned} g_n(\lambda, t, v, p) &= (a_\lambda - n\lambda)v + n\lambda f^{(n-1)} + (a + b^2(\lambda - \underline{\lambda}))p - n(\varphi^{(n)} - \varphi^{(n-1)}) \\ &+ \alpha \frac{b^2(\lambda - \underline{\lambda})v^2 + b^2(\lambda - \underline{\lambda})^2 vp + \frac{n}{2}(\varphi^{(n)} - \varphi^{(n-1)})^2 + n\lambda(\varphi^{(n)} - \varphi^{(n-1)})(v - f^{(n-1)})}{\sqrt{b^2(\lambda - \underline{\lambda})^2 v^2 + n\lambda(\varphi^{(n)} - \varphi^{(n-1)})^2}}. \end{aligned} \quad (4.7)$$

From Walter (1970, Section 28, pages 213-215), we see that we only need to verify that (3.2) holds for $v > 0 = w = q$. It is not difficult to show that Lemma 4.3 implies that g_n satisfies the inequality in (3.13). Thus, by Assumption 3.4, g_n satisfies (3.2) with the corresponding c and d satisfying the growth conditions in (3.3).

Next, note that because $f^{(n)} = \varphi_\lambda^{(n)}$ satisfies (4.5), $\mathcal{L}f^{(n)} = 0$. Also, we have $\mathcal{L}\mathbf{0} = \left(n\lambda - \alpha\sqrt{n\lambda}\right)f^{(n-1)} - \left(n - \alpha/2\sqrt{n\lambda}\right)(\varphi^{(n)} - \varphi^{(n-1)}) \leq 0$ by the induction assumption, by Lemma 4.3, and because $\lambda \geq \underline{\lambda} \geq \alpha^2$. These observations, together with $f^{(n)}(\lambda, T) = 0$, imply that $f^{(n)} = \varphi_\lambda^{(n)} \leq 0$ on G . \square

4.2. Comparative Statics for $P^{(n)}$

In this section, we present properties of $P^{(n)}$ to parallel those in Section 3.2. We show that as we vary the model parameters, the price $P^{(n)}$ responds consistently with what we

expect, as did $P = P^{(1)}$ in Section 3.2. We begin with a lemma that will help in proving Theorem 4.6 below.

Lemma 4.5. *Suppose $A \geq B$ and C are constants; then, $\sqrt{C^2 + A^2} \leq (A - B) + \sqrt{C^2 + B^2}$.*

Proof. By squaring both sides, we see that this inequality is equivalent to $C^2 + A^2 \leq (A^2 - 2AB + B^2) + (C^2 + B^2) + 2(A - B)\sqrt{C^2 + B^2}$, which simplifies to $0 \leq (A - B)(\sqrt{C^2 + B^2} - B)$, which is clearly true. \square

Theorem 4.6. *Suppose $0 \leq \alpha_1 < \alpha_2 \leq \sqrt{\lambda}$, and let $P^{(n),\alpha_i}$ be the solution to (4.1) with $\alpha = \alpha_i$, for $i = 1, 2$ and for $n \geq 0$. Then, $P^{(n),\alpha_1}(r, \lambda, t) \leq P^{(n),\alpha_2}(r, \lambda, t)$ for all $(r, \lambda, t) \in \mathbf{R}^+ \times G$.*

Proof. It is enough to show that $\varphi^{(n),\alpha_1} \leq \varphi^{(n),\alpha_2}$ on G , in which $\varphi^{(n),\alpha_i}$ has the obvious meaning. We proceed by induction. It is clear that the inequality holds for $n = 0$ because $\varphi^{0,\alpha_i} \equiv 0$ for $i = 1, 2$. For $n \geq 1$, assume that $\varphi^{(n-1),\alpha_1} \leq \varphi^{(n-1),\alpha_2}$, and show that $\varphi^{(n),\alpha_1} \leq \varphi^{(n),\alpha_2}$.

Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with $g = g_n$ from (4.3) with $\alpha = \alpha_1$. Because $\varphi^{(n),\alpha_1}$ solves (4.2) with $\alpha = \alpha_1$, we have $\mathcal{L}\varphi^{(n),\alpha_1} = 0$. Also,

$$\begin{aligned}
\mathcal{L}\varphi^{(n),\alpha_2} &= -n\lambda \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_1} \right) + n\lambda \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_2} \right) \\
&\quad + \alpha_1 \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\varphi_{\lambda}^{(n),\alpha_2} \right)^2 + n\lambda \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_1} \right)^2} \\
&\quad - \alpha_2 \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\varphi_{\lambda}^{(n),\alpha_2} \right)^2 + n\lambda \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_2} \right)^2} \\
&= -n\lambda \left(\varphi^{(n-1),\alpha_2} - \varphi^{(n-1),\alpha_1} \right) \\
&\quad + \alpha_1 \left\{ \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\varphi_{\lambda}^{(n),\alpha_2} \right)^2 + n\lambda \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_1} \right)^2} \right. \\
&\quad \quad \left. - \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\varphi_{\lambda}^{(n),\alpha_2} \right)^2 + n\lambda \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_2} \right)^2} \right\} \\
&\quad - (\alpha_2 - \alpha_1) \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\varphi_{\lambda}^{(n),\alpha_2} \right)^2 + n\lambda \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_2} \right)^2} \\
&\leq - \left(n\lambda - \alpha_1 \sqrt{n\lambda} \right) \left(\varphi^{(n-1),\alpha_2} - \varphi^{(n-1),\alpha_1} \right) \leq 0 = \mathcal{L}\varphi^{(n),\alpha_1}.
\end{aligned} \tag{4.8}$$

Note that the first inequality follows from Lemma 4.5 with $A = \sqrt{n\lambda} \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_1} \right)$, $B = \sqrt{n\lambda} \left(\varphi^{(n),\alpha_2} - \varphi^{(n-1),\alpha_2} \right)$, and $C = b(\lambda - \underline{\lambda})\varphi_{\lambda}^{(n),\alpha_2}$, from the induction hypothesis, and from $\alpha_2 > \alpha_1$. In addition, both $\varphi^{(n),\alpha_1}$ and $\varphi^{(n),\alpha_2}$ satisfy the terminal condition

$\varphi^{(n),\alpha_i}(\lambda, T) = 1$. Thus, Theorem 3.2 and Lemma 4.1 imply that $\varphi^{(n),\alpha_1} \leq \varphi^{(n),\alpha_2}$ on G . \square

Theorem 4.6 extends Theorem 3.8 and states that as the parameter α increases, the risk-adjusted price $P^{(n),\alpha}$ increases. We have the following corollary to Theorem 4.6.

Corollary 4.7. *Let $P^{(n),\alpha_0}$ be the solution to (4.1) with $\alpha = 0$; then, $P^{(n),\alpha_0} \leq P^{(n),\alpha}$ for all $0 \leq \alpha \leq \sqrt{\lambda}$, and we can express the lower bound $P^{(n),\alpha_0}$ as follows: $P^{(n),\alpha_0}(r, \lambda, t) = nF(r, t)\varphi^{\alpha_0}(\lambda, t)$, in which φ^{α_0} is given by (3.15).*

Proof. It is straightforward to show that $n\varphi^{\alpha_0}$ solves (4.2) with $\alpha = 0$, and the result follows. \square

Note that $n\varphi^{\alpha_0}$ is the expected number of survivors under the physical measure, so the lower bound of $\frac{1}{n}P^{(n)}$ (as α approaches zero) is the same as the lower bound of P , namely, $F\varphi^{\alpha_0}$.

Next, we examine how the risk-adjusted price $P^{(n)}$ varies with the drift and volatility of the stochastic hazard rate. We state the following two theorems without proof because their proofs extend those of Theorems 3.10 and 3.11, respectively, as the proof of Theorem 4.6 extends the one of Theorem 3.8.

Theorem 4.8. *Suppose $a_1(\lambda, t) \leq a_2(\lambda, t)$ on G , and let $P^{(n),a_i}$ denote the solution to (4.1) with $a = a_i$, for $i = 1, 2$. Then, $P^{(n),a_1}(r, \lambda, t) \geq P^{(n),a_2}(r, \lambda, t)$ for all $(r, \lambda, t) \in \mathbf{R}^+ \times G$.*

Theorem 4.9. *Suppose $0 \leq b_1(t) \leq b_2(t)$ on $[0, T]$, and let $P^{(n),b_i}$ denote the solution to (4.1) with $b = b_i$, for $i = 1, 2$. If $P_{\lambda\lambda}^{(n),b_i} \geq 0$ for $i = 1$ or 2 , then $P^{(n),b_1}(r, \lambda, t) \leq P^{(n),b_2}(r, \lambda, t)$ for all $(r, \lambda, t) \in \mathbf{R}^+ \times G$.*

4.3. Subadditivity of $P^{(n)}$

Subadditivity holds for the standard deviation premium principle, that is, $H(X) + H(Y) \geq H(X + Y)$; thus, we expect it to hold for our pricing rule. We next show that $P^{(n)}$ is subadditive. Specifically, we show that for m, n nonnegative integers, the following inequality holds:

$$P^{(m)} + P^{(n)} \geq P^{(m+n)}. \quad (4.9)$$

Subadditivity is a reasonable property because if it did not hold, then buyers of insurance could insure risks separately and thereby save money.

We begin with a lemma that we will help us prove (4.9).

Lemma 4.10. Suppose $A \geq C \geq B$, B_λ , and C_λ are constants; then, for nonnegative integers m and n ,

$$\sqrt{(B_\lambda + C_\lambda)^2 + (m+n)A^2} - \sqrt{n}(A-C) \leq \sqrt{B_\lambda^2 + mB^2} + \sqrt{C_\lambda^2 + nC^2} + \sqrt{m}(A-B). \quad (4.10)$$

Proof. The left-hand side of (4.10) is nonnegative because $A \geq A-C \geq 0$. Square both sides of (4.10) to get

$$\begin{aligned} & (B_\lambda + C_\lambda)^2 + (m+n)A^2 + n(A-C)^2 - 2\sqrt{n}(A-C)\sqrt{(B_\lambda + C_\lambda)^2 + (m+n)A^2} \\ & \leq B_\lambda^2 + mB^2 + C_\lambda^2 + nC^2 + m(A-B)^2 + 2\sqrt{m}(A-B) \left\{ \sqrt{B_\lambda^2 + mB^2} + \sqrt{C_\lambda^2 + nC^2} \right\} \\ & \quad + 2\sqrt{B_\lambda^2 + mB^2}\sqrt{C_\lambda^2 + nC^2}, \end{aligned} \quad (4.11)$$

which simplifies to

$$\begin{aligned} & B_\lambda C_\lambda + nA(A-C) + mB(A-B) \leq \sqrt{n}(A-C)\sqrt{(B_\lambda + C_\lambda)^2 + (m+n)A^2} \\ & \quad + \sqrt{m}(A-B) \left\{ \sqrt{B_\lambda^2 + mB^2} + \sqrt{C_\lambda^2 + nC^2} \right\} + \sqrt{B_\lambda^2 + mB^2}\sqrt{C_\lambda^2 + nC^2}. \end{aligned} \quad (4.12)$$

If the left-hand side of (4.12) is nonpositive, then we are done. Suppose the left-hand side of (4.12) is positive, so that after squaring both sides, this inequality is equivalent to

$$\begin{aligned} & B_\lambda^2 C_\lambda^2 + 2B_\lambda C_\lambda \{nA(A-C) + mB(A-B)\} + n^2 A^2 (A-C)^2 \\ & \quad + 2mnAB(A-B)(A-C) + m^2 B^2 (A-B)^2 \\ & \leq n(A-C)^2 \{(B_\lambda + C_\lambda)^2 + (m+n)A^2\} + m(A-B)^2 \{B_\lambda^2 + C_\lambda^2 + mB^2 + nC^2\} \\ & \quad + (B_\lambda^2 + mB^2)(C_\lambda^2 + nC^2) + D, \end{aligned} \quad (4.13)$$

in which $D \geq 0$ is a sum of nonnegative square-root terms. Inequality (4.13) simplifies to

$$\begin{aligned} 0 & \leq m \{C_\lambda^2 (A-B)^2 + (B_\lambda(A-B) - C_\lambda B)^2\} + n \{B_\lambda^2 (A-C)^2 + (B_\lambda C - C_\lambda(A-C))^2\} \\ & \quad + D + mn \{(A(A-C) - C(A-B))^2 + B^2 C^2 + 2A(C-B)(A-B)(A-C)\}, \end{aligned} \quad (4.14)$$

which is true because $C \geq B$. \square

Theorem 4.11. If m and n are nonnegative integers, then $P^{(m)} + P^{(n)} \geq P^{(m+n)}$.

Proof. We prove this inequality by induction on the sum $m + n$. We know that (4.9) holds when m or n equals 0 or when $m + n$ equals 0 or 1 because $P^{(0)} \equiv 0$. For $m + n \geq 2$ with $m \geq 1$ and $n \geq 1$, assume that $P^{(k)} + P^{(\ell)} \geq P^{(k+\ell)}$ for all nonnegative integers k and ℓ such that $k + \ell \leq m + n - 1$. We proceed to show that $P^{(m)} + P^{(n)} \geq P^{(m+n)}$.

The function $\xi = \varphi^{(m)} + \varphi^{(n)}$ solves

$$\begin{cases} \xi_t + a\xi_\lambda + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\xi_{\lambda\lambda} - m\lambda(\varphi^{(m)} - \varphi^{(m-1)}) - n\lambda(\varphi^{(n)} - \varphi^{(n-1)}) \\ \quad = -\alpha\sqrt{b^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{(m)})^2 + m\lambda(\varphi^{(m)} - \varphi^{(m-1)})^2} \\ \quad \quad - \alpha\sqrt{b^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{(n)})^2 + n\lambda(\varphi^{(n)} - \varphi^{(n-1)})^2}, \\ \xi(\lambda, T) = m + n, \end{cases} \quad (4.15)$$

and $\phi = \varphi^{(m+n)}$ solves

$$\begin{cases} \phi_t + a\phi_\lambda + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\phi_{\lambda\lambda} - (m+n)\lambda(\phi - \varphi^{(m+n-1)}) \\ \quad = -\alpha\sqrt{b^2(\lambda - \underline{\lambda})^2\phi_\lambda^2 + (m+n)\lambda(\phi - \varphi^{(m+n-1)})^2}, \\ \phi(\lambda, T) = m + n. \end{cases} \quad (4.16)$$

Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with $g = g_{m+n}$ from (4.3). From (4.16), we deduce that $\mathcal{L}\phi = 0$. Also, because ξ solves (4.15), we have

$$\begin{aligned} \mathcal{L}\xi &= m\lambda(\varphi^{(m)} - \varphi^{(m-1)}) + n\lambda(\varphi^{(n)} - \varphi^{(n-1)}) - (m+n)\lambda(\xi - \varphi^{(m+n-1)}) \\ &\quad - \alpha\sqrt{b^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{(m)})^2 + m\lambda(\varphi^{(m)} - \varphi^{(m-1)})^2} \\ &\quad - \alpha\sqrt{b^2(\lambda - \underline{\lambda})^2(\varphi_\lambda^{(n)})^2 + n\lambda(\varphi^{(n)} - \varphi^{(n-1)})^2} \\ &\quad + \alpha\sqrt{b^2(\lambda - \underline{\lambda})^2\xi_\lambda^2 + (m+n)\lambda(\xi - \varphi^{(m+n-1)})^2} \\ &\leq -(\varphi^{(m)} + \varphi^{(n-1)} - \varphi^{(m+n-1)}) (n\lambda - \alpha\sqrt{n\lambda}) \\ &\quad - (\varphi^{(m-1)} + \varphi^{(n)} - \varphi^{(m+n-1)}) (m\lambda - \alpha\sqrt{m\lambda}) \\ &\leq 0 = \mathcal{L}\phi. \end{aligned} \quad (4.17)$$

The last inequality follows from the induction hypothesis and from $\lambda - \alpha\sqrt{\lambda} > 0$ for $\lambda > \underline{\lambda} \geq \alpha^2$. The first inequality in (4.17) follows from Lemma 4.10 with the assignments

$B_\lambda = b(\lambda - \underline{\lambda})\varphi_\lambda^{(m)}$, $C_\lambda = b(\lambda - \underline{\lambda})\varphi_\lambda^{(n)}$, $A = \sqrt{\lambda}(\xi - \varphi^{(m+n-1)})$, $B = \sqrt{\lambda}(\varphi^{(m)} - \varphi^{(m-1)})$, and $C = \sqrt{\lambda}(\varphi^{(n)} - \varphi^{(n-1)})$. By the induction assumption, we have $A \geq B$ and $A \geq C$; without loss of generality, $C \geq B$. The functions ξ and ϕ satisfy the same terminal condition when $t = T$. Thus, Theorem 3.2 and Lemma 4.1 imply that $\xi \geq \phi$, or equivalently $P^{(m)} + P^{(n)} \geq P^{(m+n)}$. \square

Theorem 4.11 gives us another proof – alternative to the one in Theorem 4.2 – that $P^{(n)} \leq nF$, namely, $P^{(n)} \leq nP^{(1)} \leq nF$, in which we use Corollary 3.6 to assert that $P^{(1)} = P \leq F$.

4.4. Limiting Behavior of $P^{(n)}$

We next consider the limiting behavior of $P^{(n)}$. To motivate the results of this section, reconsider the standard deviation premium principle from (3.1) as applied to pricing n risks X_1, X_2, \dots, X_n that are identically distributed to a random variable X and conditionally independent given the random hazard rate. For concreteness, suppose that X_i is the indicator random variable of the event that individual i will be alive at time T . Then, $X_S = \sum_{i=1}^n X_i$ is the total number of survivors at time T , so that $\mathbf{E}(X_S) = n\mathbf{E}X$, and

$$\begin{aligned} \mathbf{Var}(X_S) &= \mathbf{Var}[\mathbf{E}(X_S|\mathcal{S})] + \mathbf{E}[\mathbf{Var}(X_S|\mathcal{S})] \\ &= n^2 \mathbf{Var}[\mathbf{E}(X|\mathcal{S})] + n\mathbf{E}[\mathbf{Var}(X|\mathcal{S})], \end{aligned} \quad (4.18)$$

in which \mathcal{S} is the σ -algebra generated by W^λ , the Brownian motion driving the stochastic hazard rate. Therefore, the static standard deviation premium principle (3.1) gives us

$$H(X_S) = n\mathbf{E}X + \alpha n \sqrt{\mathbf{Var}[\mathbf{E}(X|\mathcal{S})] + \frac{1}{n}\mathbf{E}[\mathbf{Var}(X|\mathcal{S})]}. \quad (4.19)$$

Note that $\frac{1}{n}H(X_S)$ decreases as n increases. Also, in the limit, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n}H(X_S) = \mathbf{E}X + \alpha \sqrt{\mathbf{Var}[\mathbf{E}(X|\mathcal{S})]}, \quad (4.20)$$

which is strictly greater than $\mathbf{E}X$ if $\alpha > 0$ and if $\mathbf{Var}[\mathbf{E}(X|\mathcal{S})] > 0$. This last inequality will hold if b is uniformly bounded below by $\kappa > 0$. Otherwise, if $b \equiv 0$, then $\mathbf{Var}[\mathbf{E}(X|\mathcal{S})] = 0$, and the right-hand side of (4.20) is simply $\mathbf{E}X$.

In this section, we show that $P^{(n)}$ behaves much as $H(X_S)$ does in (4.19) and (4.20). In Theorem 4.13, we show that the price per risk, $\frac{1}{n}P^{(n)}$, decreases as n increases; that is, by increasing the number of individuals insured, we reduce the risk per individual (as measured by the price). This result is consistent with what we expect, as inspired by

(4.19). The question answered by the results that follow Theorem 4.13 is *how far* does $\frac{1}{n}P^{(n)}$ decrease, and we obtain results in close parallel to (4.20).

We begin with a useful lemma.

Lemma 4.12. *If $n \geq 2$, and if $A \geq C \geq 0$ and B_λ are constants, then the following inequality holds*

$$\sqrt{B_\lambda^2 + \frac{1}{n}C^2} \leq \sqrt{n-2}(A-C) + \sqrt{B_\lambda^2 + \frac{1}{n-1}((n-1)C - (n-2)A)^2}. \quad (4.21)$$

Proof. By squaring both sides of (4.21), we can show that (4.21) is equivalent to

$$\begin{aligned} \frac{1}{n}C^2 &\leq (n-2)(A-C)^2 + \frac{1}{n-1}((n-1)C - (n-2)A)^2 \\ &\quad + 2\sqrt{n-2}(A-C)\sqrt{B_\lambda^2 + \frac{1}{n-1}((n-1)C - (n-2)A)^2}. \end{aligned} \quad (4.22)$$

Because B_λ is arbitrary, (4.22) is true if and only if (4.22) holds when $B_\lambda = 0$, which is equivalent to the following after taking the square root of the resulting right-hand side of (4.22):

$$0 \leq \sqrt{n-2} \left(1 - \sqrt{\frac{n-2}{n-1}} \right) (A-C) + C \left(\sqrt{n-1} - \frac{n-2}{\sqrt{n-1}} - \frac{1}{\sqrt{n}} \right). \quad (4.23)$$

Inequality (4.23) holds if $\sqrt{n-1} - \frac{n-2}{\sqrt{n-1}} - \frac{1}{\sqrt{n}} \geq 0$, which is true for $n \geq 2$. \square

Theorem 4.13. $\frac{1}{n}P^{(n)}$ decreases with respect to $n \geq 1$.

Proof. Define $\zeta^{(n)} = \frac{1}{n}\varphi^{(n)}$ for $n \geq 1$. Note that $\zeta^{(n)}$ solves

$$\begin{cases} \zeta_t^{(n)} + a\zeta_\lambda^{(n)} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\zeta_{\lambda\lambda}^{(n)} - \lambda \left(n\zeta^{(n)} - (n-1)\zeta^{(n-1)} \right) \\ \quad = -\alpha \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\zeta_\lambda^{(n)} \right)^2 + \frac{1}{n}\lambda \left(n\zeta^{(n)} - (n-1)\zeta^{(n-1)} \right)^2}, \\ \zeta^{(n)}(\lambda, T) = 1, \end{cases} \quad (4.24)$$

with $\zeta^{(1)} = \varphi^{(1)} = \varphi$.

We proceed by induction and show that $\zeta^{(n)} \leq \zeta^{(n-1)}$ for $n \geq 2$. We first show that $\zeta^{(2)} \leq \zeta^{(1)} = \varphi$. Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with $g = g_2$ given by

$$g_2(\lambda, t, v, p) = a(\lambda, t)p - \lambda(2v - \varphi) + \alpha \sqrt{b^2(\lambda - \underline{\lambda})^2 p^2 + \frac{1}{2}\lambda(2v - \varphi)^2}. \quad (4.25)$$

Clearly, g_2 satisfies (3.2) and (3.3), so we can apply Theorem 3.2. Note that because $\zeta^{(2)}$ solves (4.24) with $n = 2$, $\mathcal{L}\zeta^{(2)} = 0$. Also,

$$\mathcal{L}\varphi = \alpha \left\{ \sqrt{b^2(\lambda - \underline{\lambda})^2 \varphi_\lambda^2 + \frac{1}{2} \lambda \varphi^2} - \sqrt{b^2(\lambda - \underline{\lambda})^2 \varphi_\lambda^2 + \lambda \varphi^2} \right\} \leq 0 = \mathcal{L}\zeta^{(2)}. \quad (4.26)$$

Additionally, $\zeta^{(2)}(\lambda, T) = 1 = \varphi(\lambda, T)$; thus, $\zeta^{(2)} \leq \zeta^{(1)} = \varphi$.

Next, assume that for $n \geq 3$, $\zeta^{(n-1)} \leq \zeta^{(n-2)}$, and show that $\zeta^{(n)} \leq \zeta^{(n-1)}$. Define a differential operator \mathcal{D} on \mathcal{G} by (3.4) with $g = g_n$ given by

$$g_n(\lambda, t, v, p) = ap - \lambda \left(nv - (n-1)\zeta^{(n-1)} \right) + \alpha \sqrt{b^2(\lambda - \underline{\lambda})^2 p^2 + \frac{1}{n} \lambda \left(nv - (n-1)\zeta^{(n-1)} \right)^2}. \quad (4.27)$$

Clearly, g_n satisfies (3.2) and (3.3), so we can apply Theorem 3.2. Note that because $\zeta^{(n)}$ solves (4.24), $\mathcal{D}\zeta^{(n)} = 0$. Also,

$$\begin{aligned} \mathcal{D}\zeta^{(n-1)} &= (n-2)\lambda \left(\zeta^{(n-1)} - \zeta^{(n-2)} \right) \\ &\quad - \alpha \left\{ \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\zeta_\lambda^{(n-1)} \right)^2 + \frac{1}{n-1} \lambda \left((n-1)\zeta^{(n-1)} - (n-2)\zeta^{(n-2)} \right)^2} \right. \\ &\quad \left. - \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\zeta_\lambda^{(n-1)} \right)^2 + \frac{1}{n} \lambda \left(\zeta^{(n-1)} \right)^2} \right\} \\ &\leq \left((n-2)\lambda - \alpha \sqrt{(n-2)\lambda} \right) \left(\zeta^{(n-1)} - \zeta^{(n-2)} \right) \leq 0 = \mathcal{D}\zeta^{(n)}. \end{aligned} \quad (4.28)$$

The first inequality in (4.28) follows from Lemma 4.12 under the assignments $A = \sqrt{\lambda}\zeta^{(n-2)}$, $C = \sqrt{\lambda}\zeta^{(n-1)}$, and $B_\lambda = b(\lambda - \underline{\lambda})\zeta_\lambda^{(n-1)}$. We also use the induction assumption. Additionally, $\zeta^{(n)}(\lambda, T) = 1 = \zeta^{(n-1)}(\lambda, T)$; thus, $\zeta^{(n)} \leq \zeta^{(n-1)}$ on G . \square

In the next two theorems, we answer the question motivated by Theorem 4.13, that is, we determine the limiting value of the decreasing sequence $\frac{1}{n}P^{(n)}$. First, we show that $\frac{1}{n}P^{(n)}$ is bounded below by $F\beta$, in which β solves

$$\begin{cases} \beta_t + (a - \alpha b(\lambda - \underline{\lambda}))\beta_\lambda + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\beta_{\lambda\lambda} - \lambda\beta = 0, \\ \beta(\lambda, T) = 1. \end{cases} \quad (4.29)$$

Intuitively, the function β is less than φ because we replaced the square root in φ 's pde with the square root of the first term. Later, we show that $\frac{1}{n}P^{(n)}$ equals $F\beta$ in the limit. In other words, $\lim_{n \rightarrow \infty} \frac{1}{n}\varphi^{(n)} = \beta$.

In order to prove that $\frac{1}{n}P^{(n)} \geq F\beta$, we require the following lemma.

Lemma 4.14. *The function β given by (4.29) is nonincreasing with respect to λ .*

Proof. Differentiate β 's equation with respect to λ to get an equation for $f = \beta_\lambda$. The function f solves

$$\begin{cases} f_t + (a_\lambda - \alpha b - \lambda)f + (a - \alpha b(\lambda - \underline{\lambda}) + b^2(\lambda - \underline{\lambda}))f_\lambda + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2 f_{\lambda\lambda} - \beta = 0, \\ f(\lambda, T) = 0. \end{cases} \quad (4.30)$$

Thus, f 's equation is of the form in (3.4) with g given by

$$g(\lambda, t, v, p) = (a_\lambda - \alpha b - \lambda)v + (a - \alpha b(\lambda - \underline{\lambda}) + b^2(\lambda - \underline{\lambda}))p - \beta. \quad (4.31)$$

Because of the growth condition assumed for a_λ in Assumption 3.4, it is straightforward to show that g satisfies the one-sided Lipschitz condition in (3.2) with growth conditions in (3.3). Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with g given by (4.31). Because f solves (4.30), we have $\mathcal{L}f = 0$. Denote by $\mathbf{0}$ the function that is identically 0 on G ; then, $\mathcal{L}\mathbf{0} = -\beta \leq 0 = \mathcal{L}f$. In addition, $f(\lambda, T) = 0$; thus, Theorem 3.2 implies that $f = \beta_\lambda \leq 0$. \square

Theorem 4.15.

$$\lim_{n \rightarrow \infty} \frac{1}{n}P^{(n)}(r, \lambda, t) \geq F(r, t)\beta(\lambda, t). \quad (4.32)$$

Proof. It is enough to show that the solution β of (4.29) is a lower bound of $\frac{1}{n}\varphi^{(n)}$. We proceed by induction. First, show that $\beta \leq \varphi = \varphi^{(1)}$. Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with g given by (3.8). Because φ solves (2.18), $\mathcal{L}\varphi = 0$. Also,

$$\mathcal{L}\beta = \alpha \left\{ \sqrt{b^2(\lambda - \underline{\lambda})^2 \beta_\lambda^2 + \lambda \beta^2} - b(\lambda - \underline{\lambda})|\beta_\lambda| \right\} \geq 0 = \mathcal{L}\varphi. \quad (4.33)$$

In addition, $\beta(\lambda, T) = \varphi(\lambda, T) = 1$; thus, Theorem 3.2 and Lemma 3.3 imply that $\beta \leq \varphi$.

Next, assume that $\beta \leq \frac{1}{n-1}\varphi^{(n-1)}$, and show that $\beta \leq \frac{1}{n}\varphi^{(n)}$. Recall that the function $\zeta^{(n)} = \frac{1}{n}\varphi^{(n)}$ solves (4.24). Define a differential operator \mathcal{D} on \mathcal{G} by (3.4) with $g = g_n$ given in (4.27). Recall that $(n-1)\zeta^{(n-1)} = \varphi^{(n-1)}$ in (4.27). Then, $\mathcal{D}\zeta^{(n)} = 0$, and

$$\begin{aligned}
\mathcal{D}\beta &= \beta_t + a\beta_\lambda + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\beta_{\lambda\lambda} - \lambda \left(n\beta - \varphi^{(n-1)} \right) \\
&\quad + \alpha \sqrt{b^2(\lambda - \underline{\lambda})^2\beta_\lambda^2 + \frac{1}{n}\lambda \left(n\beta - \varphi^{(n-1)} \right)^2} \\
&= \lambda \left(\varphi^{(n-1)} - (n-1)\beta \right) + \alpha \left\{ \sqrt{b^2(\lambda - \underline{\lambda})^2\beta_\lambda^2 + \frac{1}{n}\lambda \left(n\beta - \varphi^{(n-1)} \right)^2} - b(\lambda - \underline{\lambda})|\beta_\lambda| \right\} \\
&\geq 0 = \mathcal{D}\zeta^{(n)}.
\end{aligned} \tag{4.34}$$

Also, $\zeta^{(n)}(\lambda, T) = \beta(\lambda, T) = 1$; thus, Theorem 3.2 and Lemma 4.1 imply that $\beta \leq \zeta^{(n)} = \frac{1}{n}\varphi^{(n)}$. \square

The next theorem tightens the result of Theorem 4.15 and shows that we have equality in (4.32). Consider the solution to the following pde:

$$\begin{cases} \gamma_t^{(n)} + (a - \alpha b(\lambda - \underline{\lambda}))\gamma_\lambda^{(n)} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\gamma_{\lambda\lambda}^{(n)} - (n\lambda - \alpha\sqrt{n\lambda}) \left(\gamma^{(n)} - \gamma^{(n-1)} \right) = 0, \\ \gamma^{(n)}(\lambda, T) = n, \end{cases} \tag{4.35}$$

in which $\gamma^{(0)} \equiv 0$. We proceed by showing in a series of lemmas that $\gamma^{(n)} \geq \varphi^{(n)}$ for $n \geq 0$. Intuitively, $\gamma^{(n)}$ is greater than $\varphi^{(n)}$ because we replaced the square root in $\varphi^{(n)}$'s pde with the sum of the square roots of the two terms. Finally, we show that $\frac{1}{n}\gamma^{(n)} - \beta$ goes to zero as n goes to infinity.

Lemma 4.16. *The function $\gamma^{(n)}$ given by (4.35) is nonincreasing with respect to λ and $\gamma^{(n+1)} \geq \gamma^{(n)}$ for $n \geq 0$.*

Proof. The proof of that $\gamma_\lambda^{(n)} \leq 0$ is similar to the proof that $\varphi_\lambda^{(n)} \leq 0$ in Theorems 3.7 and 4.4, so we omit the details. Similarly, $\gamma^{(n+1)} \geq \gamma^{(n)}$ follows as in the proof of Lemma 4.3. \square

Lemma 4.17. *$\gamma^{(n)} \geq \varphi^{(n)}$ for $n \geq 0$.*

Proof. The result is true for $n = 0$ because $\gamma^{(0)} = \varphi^{(0)} = 0$. Suppose for $n \geq 1$, we have $\gamma^{(n-1)} \geq \varphi^{(n-1)}$, and show that $\gamma^{(n)} \geq \varphi^{(n)}$. Define a differential operator \mathcal{L} on \mathcal{G} by (3.4) with $g = g_n$ given in (4.3). Then, $\mathcal{L}\varphi^{(n)} = 0$, and

$$\begin{aligned}
\mathcal{L}\gamma^{(n)} &= \alpha b(\lambda - \underline{\lambda})\gamma_{\lambda}^{(n)} - n\lambda \left(\gamma^{(n)} - \varphi^{(n-1)} \right) + \left(n\lambda - \alpha\sqrt{n\lambda} \right) \left(\gamma^{(n)} - \gamma^{(n-1)} \right) \\
&\quad + \alpha \sqrt{b^2(\lambda - \underline{\lambda})^2 \left(\gamma_{\lambda}^{(n)} \right)^2 + n\lambda \left(\gamma^{(n)} - \varphi^{(n-1)} \right)^2} \\
&\leq \left(n\lambda - \alpha\sqrt{n\lambda} \right) \left(\varphi^{(n-1)} - \gamma^{(n-1)} \right) \leq 0 = \mathcal{L}\varphi^{(n)}.
\end{aligned} \tag{4.36}$$

In the first inequality in (4.36), we use the facts that $\gamma_{\lambda}^{(n)} \leq 0$ and $\gamma^{(n)} \geq \gamma^{(n-1)}$, and we use the subadditivity of the square root function. In the second inequality, we use the induction assumption. We also have the terminal conditions $\gamma^{(n)}(\lambda, T) = n = \varphi^{(n)}(\lambda, T)$; thus, Theorem 3.2 and Lemma 4.1 imply that $\gamma^{(n)} \geq \varphi^{(n)}$. \square

Now, we are ready to prove the main theorem of this section.

Theorem 4.18.

$$\lim_{n \rightarrow \infty} \frac{1}{n} P^{(n)}(r, \lambda, t) = F(r, t) \beta(\lambda, t). \tag{4.37}$$

Proof. By Theorem 4.15 and Lemma 4.17, the theorem is proved if we show that $\frac{1}{n}\gamma^{(n)} - \beta$ goes to zero as n goes to infinity because $\frac{1}{n}\gamma^{(n)} - \beta \geq \frac{1}{n}\varphi^{(n)} - \beta \geq 0$.

Define $\Phi^{(n)}$ on G by $\Phi^{(n)} = \frac{1}{n}\gamma^{(n)} - \beta$, so the theorem is proved if we show that $\lim_{n \rightarrow \infty} \Phi^{(n)}(\lambda, t) = 0$. The function $\Phi^{(n)}$ solves the recursion

$$\begin{cases} \Phi_t^{(n)} + (a - \alpha b(\lambda - \underline{\lambda}))\Phi_{\lambda}^{(n)} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2\Phi_{\lambda\lambda}^{(n)} - \left(n\lambda - \alpha\sqrt{n\lambda} \right) \Phi^{(n)} \\ \quad = -\alpha\sqrt{\frac{\lambda}{n}}\beta - (n-1) \left(\lambda - \alpha\sqrt{\frac{\lambda}{n}} \right) \Phi^{(n-1)}, \\ \Phi^{(n)}(\lambda, T) = 0, \end{cases} \tag{4.38}$$

with $0 \leq \Phi^{(1)} = \gamma^{(1)} - \beta \leq 1$. From (4.38) and the Feynman-Kac Theorem, we deduce the following expression for $\Phi^{(n)}$ in terms of $\Phi^{(n-1)}$:

$$\begin{aligned}
\Phi^{(n)}(\lambda, t) &= \alpha \tilde{\mathbf{E}} \left[\int_t^T \sqrt{\frac{\lambda_s}{n}} \beta(\lambda_s, s) e^{-\int_t^s (n\lambda_u - \alpha\sqrt{n\lambda_u}) du} ds \middle| \lambda_t = \lambda \right] \\
&\quad + (n-1) \tilde{\mathbf{E}} \left[\int_t^T \left(\lambda_s - \alpha\sqrt{\frac{\lambda_s}{n}} \right) \Phi^{(n-1)}(\lambda_s, s) e^{-\int_t^s (n\lambda_u - \alpha\sqrt{n\lambda_u}) du} ds \middle| \lambda_t = \lambda \right],
\end{aligned} \tag{4.39}$$

in which λ follows the process $d\lambda_s = (a - \alpha b(\lambda_s - \underline{\lambda}))ds + b(\lambda_s - \underline{\lambda})d\tilde{W}_s^\lambda$, with $\tilde{W}_s^\lambda = W_s^\lambda + \alpha s$. The process \tilde{W}^λ is a standard Brownian motion with respect to the probability space $(\Omega, \mathcal{F}, \tilde{\mathbf{P}})$, in which $\frac{d\tilde{\mathbf{P}}}{d\mathbf{P}} = e^{-\alpha W_T - \frac{1}{2}\alpha^2 T}$. $\tilde{\mathbf{E}}$ denotes expectation with respect to $\tilde{\mathbf{P}}$. Note that α is analogous to the bond market's price of risk q in (2.4).

Suppose $\Phi^{(n-1)}(\lambda, t) \leq K_{n-1}$ on G for some $n \geq 2$. Then, we can bound $\Phi^{(n)}$ as follows:

$$\begin{aligned} \Phi^{(n)}(\lambda, t) &\leq \alpha \tilde{\mathbf{E}} \left[\int_t^T \sqrt{\frac{\lambda_s}{n}} e^{-\int_t^s (n\lambda_u - \alpha\sqrt{n\lambda_u})du} ds \middle| \lambda_t = \lambda \right] \\ &\quad + (n-1)K_{n-1} \tilde{\mathbf{E}} \left[\int_t^T \left(\lambda_s - \alpha\sqrt{\frac{\lambda_s}{n}} \right) e^{-\int_t^s (n\lambda_u - \alpha\sqrt{n\lambda_u})du} ds \middle| \lambda_t = \lambda \right], \end{aligned} \quad (4.40)$$

in which we use the fact that $0 \leq \beta \leq 1$. Define $f^{(n)}$ and $h^{(n)}$ for $n \geq 2$ by

$$f^{(n)}(\lambda, t) = \alpha \tilde{\mathbf{E}} \left[\int_t^T n\sqrt{\lambda_s} e^{-\int_t^s (n\lambda_u - \alpha\sqrt{n\lambda_u})du} ds \middle| \lambda_t = \lambda \right], \quad (4.41)$$

and

$$h^{(n)}(\lambda, t) = \tilde{\mathbf{E}} \left[\int_t^T \left(n\lambda_s - \alpha\sqrt{n\lambda_s} \right) e^{-\int_t^s (n\lambda_u - \alpha\sqrt{n\lambda_u})du} ds \middle| \lambda_t = \lambda \right]. \quad (4.42)$$

Thus, inequality (4.40) is equivalent to

$$\Phi^{(n)}(\lambda, t) \leq \frac{1}{n^{3/2}} f^{(n)}(\lambda, t) + \frac{n-1}{n} K_{n-1} h^{(n)}(\lambda, t). \quad (4.43)$$

After proving the following two lemmas that give us bounds on $f^{(n)}$ and $h^{(n)}$ in (4.41) and (4.42), respectively, we finish the proof of Theorem 4.18.

Lemma 4.19. *The function $f^{(n)}$ defined by (4.41) is bounded above by $J = \frac{\alpha\sqrt{2}}{\sqrt{2\lambda} - \alpha}$ for all $n \geq 2$.*

Proof. By the Feynman-Kac Theorem, $f^{(n)}$ solves the linear pde

$$\begin{cases} f_t^{(n)} + (a - \alpha b(\lambda - \underline{\lambda}))f_\lambda^{(n)} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2 f_{\lambda\lambda}^{(n)} - (n\lambda - \alpha\sqrt{n\lambda})f^{(n)} = -\alpha n\sqrt{\lambda}, \\ f^{(n)}(\lambda, T) = 0. \end{cases} \quad (4.44)$$

Define a differential operator \mathcal{L} on \mathcal{G} via $f^{(n)}$'s equation. Let \mathbf{J} denote the function that is identically equal to J . Thus, $\mathcal{L}f^{(n)} = 0$, and $\mathcal{L}\mathbf{J} = -\left(n\lambda - \alpha\sqrt{n\lambda}\right)J + \alpha n\sqrt{\lambda} \leq 0 = \mathcal{L}f^{(n)}$ for $n \geq 2$. We also have $f^{(n)}(\lambda, T) = 0 \leq J$; thus, from Theorem 3.2, we conclude that $f^{(n)} \leq J$ on G for $n \geq 2$. \square

Lemma 4.20. *The function $h^{(n)}$ defined by (4.42) is bounded above by 1 for all $n \geq 2$.*

Proof. By the Feynman-Kac Theorem, $h^{(n)}$ solves the linear pde

$$\begin{cases} h_t^{(n)} + (a - \alpha b(\lambda - \underline{\lambda}))h_\lambda^{(n)} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2 h_{\lambda\lambda}^{(n)} - (n\lambda - \alpha\sqrt{n\lambda})h^{(n)} = -(n\lambda - \alpha\sqrt{n\lambda}), \\ h^{(n)}(\lambda, T) = 0. \end{cases} \quad (4.45)$$

Define a differential operator \mathcal{L} on \mathcal{G} via $h^{(n)}$'s equation; thus, $\mathcal{L}h^{(n)} = 0 = \mathcal{L}\mathbf{1}$. We also have $h^{(n)}(\lambda, T) = 0 \leq 1$; thus, from Theorem 3.2, we conclude that $h^{(n)} \leq 1$ on G for $n \geq 2$. \square

End of Proof of Theorem 4.18. We have shown that if $\Phi^{(n-1)} \leq K_{n-1}$, then $\Phi^{(n)} \leq K_n$, in which for $n \geq 2$,

$$K_n = \frac{J}{n^{3/2}} + \frac{n-1}{n}K_{n-1}, \quad (4.46)$$

with $K_1 = 1$. In hindsight, we could have used induction to show that if $\Phi^{(n-1)} \leq K_{n-1}$, then $\Phi^{(n)} \leq K_n$ via a comparison argument with a differential operator based on the pde in (4.38). However, we have chosen to leave the proof as is because of its constructive nature in demonstrating the origin of (4.46).

Define $L_n = nK_n$; thus,

$$L_n = L_{n-1} + \frac{J}{\sqrt{n}}, \quad n \geq 2, \quad (4.47)$$

from which it follows that

$$L_n = 1 + \sum_{i=2}^n \frac{J}{\sqrt{i}} \leq 1 + J \int_1^n \frac{dx}{\sqrt{x}} \leq 1 + 2J\sqrt{n}, \quad n \geq 2. \quad (4.48)$$

Finally, we have

$$\Phi^{(n)}(\lambda, t) \leq K_n \leq \frac{1}{n} + \frac{2J}{\sqrt{n}}, \quad n \geq 1, \quad (4.49)$$

The right-hand side of inequality (4.49) goes to zero as n goes to ∞ ; thus, $\Phi^{(n)}(\lambda, t)$ goes to zero as n goes to ∞ . In other words, $\lim_{n \rightarrow \infty} \frac{1}{n} P^{(n)}(\lambda, r, t) = F(r, t) \beta(\lambda, t)$, as we wished to show. \square

We have the following corollaries of Theorem 4.18. The first parallels the observation that the right-hand side in (4.20) is simply $\mathbf{E}X$ if $b \equiv 0$.

Corollary 4.21. *If $b \equiv 0$, then $\lim_{n \rightarrow \infty} \frac{1}{n} P^{(n)}(\lambda, r, t) = F(r, t) \varphi^{\alpha 0}$ in which $\varphi^{\alpha 0}$ is the physical probability of survival given in (3.15).*

Proof. From Theorem 4.18, we know that $\frac{1}{n} \varphi^{(n)}$ goes to β as n goes to infinity. Therefore, the corollary follows because $\beta = \varphi^{\alpha 0}$ when $b \equiv 0$, which is clear from (4.29) by setting b equal to 0. \square

Corollary 4.22. *If b is uniformly bounded below by $\kappa > 0$, then $\lim_{n \rightarrow \infty} \frac{1}{n} P^{(n)}(\lambda, r, t) \geq F(r, t) \varphi^{\alpha 0}$, with equality only when $t = T$.*

Proof. This result follows from the fact that $\beta \geq \varphi^{\alpha 0}$, with equality only when $t = T$. Indeed, define a differential operator \mathcal{L} on \mathcal{G} by (3.4) and (3.8) with $\alpha = 0$. Thus, $\mathcal{L}\varphi^{\alpha 0} = 0$, and

$$\mathcal{L}\beta = \alpha b(\lambda - \underline{\lambda})\beta_{\lambda} \leq 0 = \mathcal{L}\varphi^{\alpha 0}. \quad (4.50)$$

Also, $\varphi^{\alpha 0}(\lambda, T) = \beta(\lambda, T) = 1$; thus, Theorem 3.2 and Lemma 3.3 imply that $\varphi^{\alpha 0} \leq \beta$.

To show that $\varphi^{\alpha 0} < \beta$ for $t \in [0, T)$, first consider the pde of $\varphi_{\lambda}^{\alpha 0}$. The function $f = \varphi_{\lambda}^{\alpha 0}$ solves

$$\begin{cases} f_t + (a_{\lambda} - \lambda)f + (a + b^2(\lambda - \underline{\lambda}))f_{\lambda} + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2 f_{\lambda\lambda} - \varphi^{\alpha 0} = 0, \\ f(\lambda, T) = 0. \end{cases} \quad (4.51)$$

From the linear pde in (4.51) and the Feynman-Kac Theorem, we deduce that

$$\varphi_{\lambda}^{\alpha 0}(\lambda, t) = -\hat{\mathbf{E}} \left[\int_t^T \varphi^{\alpha 0}(\lambda_s, s) e^{-\int_t^s (\lambda_u - a_{\lambda}(\lambda_u, u)) du} ds \middle| \lambda_s = \lambda \right], \quad (4.52)$$

in which λ follows the diffusion $d\lambda_s = (a + b^2(\lambda_s - \underline{\lambda}))ds + b(\lambda_s - \underline{\lambda})d\hat{W}_s^{\lambda}$, with $\hat{W}_s^{\lambda} = W_s^{\lambda} - \int_0^s b(u)du$. The process \hat{W}^{λ} is a standard Brownian motion with respect to a suitably-defined probability space, and $\hat{\mathbf{E}}$ denotes expectation on that space. Note that $\varphi_{\lambda}^{\alpha 0}(\lambda, t) < 0$ for $\lambda > \underline{\lambda}$ and for $t \in [0, T)$ because $\varphi^{\alpha 0}(\lambda, t) > 0$ on that domain by the representation in (3.15).

Next, consider the pde of $B = \beta - \varphi^{\alpha 0}$.

$$\begin{cases} B_t + (a - \alpha b(\lambda - \underline{\lambda}))B_\lambda + \frac{1}{2}b^2(\lambda - \underline{\lambda})^2 B_{\lambda\lambda} - \lambda B = \alpha b(\lambda - \underline{\lambda})\varphi_\lambda^{\alpha 0}, \\ B(\lambda, T) = 0. \end{cases} \quad (4.53)$$

From the linear pde in (4.53) and the Feynman-Kac Theorem, we deduce that

$$B(\lambda, t) = -\alpha \tilde{\mathbf{E}} \left[\int_t^T b(s)(\lambda_s - \underline{\lambda})\varphi_\lambda^{\alpha 0}(\lambda_s, s) e^{-\int_t^s \lambda_u du} ds \middle| \lambda_s = \lambda \right], \quad (4.54)$$

in which λ follows the diffusion from the paragraph following (4.39) and $\tilde{\mathbf{E}}$ is as in (4.39). Note that $B(\lambda, t) > 0$ for $\lambda > \underline{\lambda}$ and for $t \in [0, T)$ because $\varphi_\lambda^{\alpha 0}(\lambda, t) < 0$ on that domain by the representation in (4.52). \square

After all this work, we can finally decompose the risk charge $P - P^{\alpha 0}$ – first mentioned in the Introduction, then briefly in Section 2.2, and again following Corollary 3.9 – into its component risk charges: one for a finite portfolio and another for stochastic mortality. More generally, we decompose the per-risk risk charge $\frac{1}{n}P^{(n)} - P^{\alpha 0}$ when the insurer holds a portfolio of n risks. Recall from Theorem 4.18 that $\lim_{n \rightarrow \infty} \frac{1}{n}P^{(n)} = F\beta$. Therefore, define $\frac{1}{n}P^{(n)} - F\beta$ as the risk charge (per risk) for holding a finite portfolio, and define $F\beta - P^{\alpha 0}$ as the risk charge for stochastic mortality even after selling to an arbitrarily large group. Thus, we have

$$\frac{1}{n}P^{(n)} - P^{\alpha 0} = \left(\frac{1}{n}P^{(n)} - F\beta \right) + (F\beta - P^{\alpha 0}) = F \left(\frac{1}{n}\varphi^{(n)} - \beta \right) + F(\beta - \varphi^{\alpha 0}), \quad (4.55)$$

in which the risk charge for stochastic mortality, namely $F(\beta - \varphi^{\alpha 0})$, is zero if $b \equiv 0$ by Corollary 4.21 and is positive (for $t < T$) if $b \geq \kappa > 0$ by Corollary 4.22.

5. Summary and Conclusions

We developed a theoretical foundation for valuing mortality risk by assuming that the risk is “priced” via the instantaneous Sharpe ratio. Because the market for pure endowments is incomplete, one cannot assert that there is a unique price. However, we believe that the price that our method produces is a valid one because of the many desirable properties that it satisfies. In particular, we studied properties of the price for n conditionally independent and identically distributed pure endowment risks. In Theorem 4.11, we showed that the price is subadditive with respect to n , and in Theorem 4.13, we

showed that the risk charge per person decreases as n increases. We also proved that if the hazard rate is deterministic, then the risk charge per person goes to zero as n goes to infinity (Theorem 4.20 and Corollary 4.21). Moreover, we proved that if the hazard rate is stochastic, then the risk charge per person is positive as n goes to infinity, which reflects the fact that the mortality risk is not diversifiable in this case (Theorem 4.20 and Corollary 4.22). Additionally, in equation (4.55), we decomposed the per-risk risk charge into the finite portfolio and stochastic mortality risk charges. Because of these properties, we anticipate that our pricing methodology will prove useful in pricing risks in other incomplete markets.

In addition to addressing the problem of the breakdown in the law of large numbers, the study of such a mortality risk premium might breathe new life into the analysis of exotic options that are embedded within insurance and pension contracts, which have traditionally been viewed as being out-of-the-money and, hence, valueless. Indeed, our theoretical framework – which incorporates a risk measure for mortality risk – could be used to evaluate the risks inherent in a number of recent industry trends, namely (i) the emergence of longevity-linked bonds, (ii) capacity constraints in the immediate annuity market, and (iii) the decline of defined benefit pension plans.

Acknowledgements

We thank Erhan Bayraktar, Kristen S. Moore, Jeffrey Rauch, Keith Promislow, and Zhengfang Zhou for their valuable help.

References

- Biffis, E. (2005), Affine processes for dynamic mortality and actuarial valuation, *Insurance: Mathematics and Economics*, to appear.
- Björk, T. (1998), *Arbitrage Theory in Continuous Time*, Oxford University Press, Oxford.
- Blanchet-Scalliet, C., N. El Karoui, and L. Martellini (2005), Dynamic asset pricing theory with uncertain time-horizon, *Journal of Economic Dynamics and Control*, 29: 1737-1764.
- Boyle, P. P. and M. Hardy (2003), Guaranteed annuity options, *ASTIN Bulletin*, 33: 125-152.
- Brennan, M. J. and E. Schwartz (1976), The pricing of equity-linked life insurance policies with an asset value guarantee, *Journal of Financial Economics*, 3 (1): 195-213.
- Cairns, A. J. G., D. Blake and K. Dowd (2004), Pricing framework for securitization of mortality risk, working paper, Heriot-Watt University.

- Cox, S. H. and Y. Lin (2004), Natural hedging of life and annuity mortality risks, *Journal of Risk and Insurance*, to appear.
- Dahl, M. (2004), Stochastic mortality in life insurance: Market reserves and mortality-linked insurance contracts, *Insurance: Mathematics and Economics*, 35: 113-136.
- DiLorenzo, E. and M. Sibillo (2003), Longevity risk: Measurement and application perspectives, working paper, Universita degli Studi di Napoli.
- Gerber, H. U. (1979), *Introduction to Mathematical Risk Theory*, Huebner Foundation Monograph 8, Wharton School of the University of Pennsylvania, Richard D. Irwin, Homewood, IL.
- Gerber, H. U. and E. S. W. Shiu (1994), Option pricing by Esscher transforms (with discussions), *Transactions of the Society of Actuaries*, 46: 99-191.
- Karatzas, I. and S. E. Shreve (1991), *Brownian Motion and Stochastic Calculus*, second edition, Springer-Verlag, New York.
- Lamberton, D. and B. Lapeyre (1996), *Introduction to Stochastic Calculus Applied to Finance*, Chapman & Hall/CRC, Boca Raton, Florida.
- Lee, R. D. and L. R. Carter (1992), Modeling and forecasting U.S. mortality, *Journal of the American Statistical Association*, 87 (419): 659-671.
- Milevsky, M. A. and S. D. Promislow (2001), Mortality derivatives and the option to annuitize, *Insurance: Mathematics and Economics*, 29: 299-318.
- Norberg, R. (2004), Vasicek beyond the normal, *Mathematical Finance*, 14 (4): 585-604.
- Olivieri, A. (2001), Uncertainty in mortality projections: An actuarial perspective, *Insurance: Mathematics and Economics*, 29: 231-245.
- Protter, P. (1995), *Stochastic Integration and Differential Equations*, Applications in Mathematics, 21, Springer-Verlag, Berlin.
- Royden, H. L. (1968), *Real Analysis*, second edition, Macmillan, New York.
- Schrager, D.F. (2005), Affine stochastic mortality, *Insurance: Mathematics and Economics*, to appear.
- Smith, A., I. Moran, and D. Walczak (2003), Why can financial firms charge for diversifiable risk?, working paper, Deloitte Touche Tohmatsu.
- Soininen, P. (1995), Stochastic variation of interest and mortality, *Proceedings of the 5th AFIR International Colloquium*, 871-904.
- Walter, W. (1970), *Differential and Integral Inequalities*, Springer-Verlag, New York.

- Windcliff, H., J. Wang, P. A. Forsyth, and K. R. Vetzal (2005), Hedging with a correlated asset: Solution of a nonlinear pricing pde, working paper, University of Waterloo.
- Zariphopoulou, T., (2001), Stochastic control methods in asset pricing, *Handbook of Stochastic Analysis and Applications*, D. Kannan and V. Lakshmikantham (editors), Marcel Dekker, New York.